

STIC-Biotech/ChemLib

From: Fredman, Jeffrey
Sent: Wednesday, February 11, 2004 11:07 AM
To: STIC-Biotech/ChemLib
Cc: Chernyshev, Olga
Subject: FW: sequence search request

PLEASE RUSH.

I Approve this large search.

Jeff Fredman

-----Original Message-----

From: Chernyshev, Olga
Sent: Wednesday, February 11, 2004 10:12 AM
To: Fredman, Jeffrey
Subject: sequence search request

Please authorize the following sequence search request:

Please search US case 09/901,187 SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, 8, 9,10, 11 and 12 in regular and pending databases. Only direct hits, because they claim these short peptides precisely, no open language.
Thank you very much!

Olga N. Chernyshev
AU1646
REM 4E84
20870
mail box 4D70

4C70

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 12
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/11/04
Date Completed: 2/12/04
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: QSP
WWW/Internet: _____
Other (specify): _____

RECEIVED
FEB 11 2004
STIC-BIOTECH/CHEM.LIB (STIC)

174119

Publication No. US20030124679A1
 GENERAL INFORMATION:
 APPLICANT: Short, Jay M.
 APPLICANT: Parakkar, Ashish
 APPLICANT: Varoglu, Mustafa
 APPLICANT: Mathur, Eric J.
 TITLE OF INVENTION: GLYCOSYLATED KINAMYCINS AND METHODS OF
 TITLE OF INVENTION: MAKING AND USING THEM
 FILE REFERENCE: 09010-280001
 CURRENT APPLICATION NUMBER: US/10/187,267A
 CURRENT FILING DATE: 2003-01-27
 PRIOR APPLICATION NUMBER: US 60/301,401
 PRIOR FILING DATE: 2001-06-27
 NUMBER OF SEQ ID NOS: 65
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 19
 LENGTH: 148
 TYPE: PRT
 ORGANISM: Streptomyces murayamaensis ATCC 21414
 US-10-187-267A-19

Query Match 97.7%; Score 43; DB 15; Length 148;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THHHPS 7
 ||:||||
 Db 93 THVHPS 99

RESULT 3
 US-09-801-944B-233
 Sequence 233, Application US/09801944B
 Publication No. US20040014169A1
 GENERAL INFORMATION:
 APPLICANT: Vogel, Gabriel
 APPLICANT: Wood, Linda S.
 TITLE OF INVENTION: No. US20040014169A1 G Protein-Coupled Receptors
 FILE REFERENCE: 00100US1
 CURRENT APPLICATION NUMBER: US/09/801,944B
 CURRENT FILING DATE: 2003-08-01
 PRIOR APPLICATION NUMBER: 60/187,828
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,715
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,929
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,930
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,825
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,833
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,830
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,829
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,582
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/187,581
 PRIOR FILING DATE: 2000-03-08
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 273
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 233
 LENGTH: 178
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-801-944B-233

Query Match 88.6%; Score 39; DB 12; Length 178;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHPHS 7
 ||:||||
 Db 102 HHHPHS 107

RESULT 4
 US-10-341-434-14
 Sequence 14, Application US/10341434
 Publication No. US20030215835A1
 GENERAL INFORMATION:
 APPLICANT: Origene Technologies
 TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
 FILE REFERENCE: 9U 204 205 R1
 CURRENT APPLICATION NUMBER: US/10/341,434
 CURRENT FILING DATE: 2003-07-18
 PRIOR APPLICATION NUMBER: US 60/348,164
 PRIOR FILING DATE: 2002-01-15
 PRIOR APPLICATION NUMBER: US 60/348,119
 PRIOR FILING DATE: 2002-01-15
 NUMBER OF SEQ ID NOS: 238
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 14
 LENGTH: 402
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-341-434-14

Query Match 88.6%; Score 39; DB 12; Length 402;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
 ||:||||
 Db 247 THCHPS 253

RESULT 5
 US-10-183-116-99
 Sequence 99, Application US/10183116
 Publication No. US20030092035A1
 GENERAL INFORMATION:
 APPLICANT: Anderson, David J.
 APPLICANT: Dong, Xinzhong
 APPLICANT: Zylka, Mark
 APPLICANT: Simon, Melvin
 APPLICANT: Han, Sang-kyou
 TITLE OF INVENTION: PAIN SIGNALING MOLECULES
 FILE REFERENCE: CALTE.4C1CP1
 CURRENT APPLICATION NUMBER: US/10/183,116
 CURRENT FILING DATE: 2002-06-26
 PRIOR APPLICATION NUMBER: US 60/222,344
 PRIOR FILING DATE: 2000-08-01
 PRIOR APPLICATION NUMBER: US 60/202,027
 PRIOR FILING DATE: 2000-05-04
 PRIOR APPLICATION NUMBER: US 09/704,707
 PRIOR FILING DATE: 2000-11-03
 PRIOR APPLICATION NUMBER: US 60/285,493
 PRIOR FILING DATE: 2001-04-19
 PRIOR APPLICATION NUMBER: US 09/849,869
 PRIOR FILING DATE: 2001-05-04
 NUMBER OF SEQ ID NOS: 109
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 99
 LENGTH: 262
 TYPE: PRT
 ORGANISM: Mus musculus
 US-10-183-116-99

Query Match 84.1%; Score 37; DB 15; Length 262;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      2 HHHPS 7
      |:|||||
Db      208 HLHPS 213

RESULT 6
US-10-183-116-91
; Sequence 91, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CP1
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-183-116-91

      Query Match      84.1%; Score 37; DB 15; Length 270;
      Best Local Similarity 83.3%; Pred. No. 1.7e+02;
      Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
      |:|||||
Db      216 HLHPS 221

RESULT 7
US-10-115-482-50
; Sequence 50, Application US/10115482
; Publication No. US20030212257A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS
; TITLE OF INVENTION: OF USING THE SAME
; FILE REFERENCE: 21404-322D
; CURRENT APPLICATION NUMBER: US/10/115,482
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/285,325
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/285,890
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286,068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286,292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/287,213
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/288,257
; PRIOR FILING DATE: 2001-05-02

      Query Match      84.1%; Score 37; DB 12; Length 911;
      Best Local Similarity 83.3%; Pred. No. 5.4e+02;
      Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
      |:|||||
Db      320 HLHPS 325

RESULT 8
US-10-115-482-48
; Sequence 48, Application US/10115482
; Publication No. US20030212257A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS
; TITLE OF INVENTION: OF USING THE SAME
; FILE REFERENCE: 21404-322D
; CURRENT APPLICATION NUMBER: US/10/115,482
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/285,325
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/285,890
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286,068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286,292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/287,213
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/288,257
; PRIOR FILING DATE: 2001-05-02
```



```
; PRIOR APPLICATION NUMBER: 60/291,134
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/291,725
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/294,771
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/296,965
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/299,128
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 149
; SEQ ID NO 48
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-482-48

Query Match      84.1%; Score 37; DB 12; Length 915;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
      |:||||
Db      320 HLHPS 325

RESULT 9
US-10-017-161-740
; Sequence 740, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 740
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-740

Query Match      84.1%; Score 37; DB 12; Length 1021;
Best Local Similarity 83.3%; Pred. No. 6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
      |:||||
Db      361 HLHPS 366

RESULT 10
US-10-335-687A-2
; Sequence 2, Application US/10335687A
; Publication No. US20030166222A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 39267, Human Kinase Family Members and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: MPI02-001PIRNM
; CURRENT APPLICATION NUMBER: US/10/335,687A
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: 60/345,773
```

```
; PRIOR FILING DATE: 2002-01-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1818
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-335-687A-2

Query Match      84.1%; Score 37; DB 12; Length 1818;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
      |:||||
Db      1223 HLHPS 1228

RESULT 11
US-10-335-687A-5
; Sequence 5, Application US/10335687A
; Publication No. US20030166222A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 39267, Human Kinase Family Members and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: MPI02-001PIRNM
; CURRENT APPLICATION NUMBER: US/10/335,687A
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: 60/345,773
; PRIOR FILING DATE: 2002-01-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1824
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-335-687A-5

Query Match      84.1%; Score 37; DB 12; Length 1824;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
      |:||||
Db      1223 HLHPS 1228

RESULT 12
US-10-082-830-180
; Sequence 180, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 180
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-830-180
```

Query Match 81.8%; Score 36; DB 15; Length 46;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THHHP 6
 :|||||
 Db 27 SHHHP 32

RESULT 13
 US-09-864-761-38417
 ; Sequence 38417, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 38417
 ; LENGTH: 49
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO Z98884.11
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.67
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
 OTHER INFORMATION: EST HUMAN HIT: BE843827.1, EVALUE 1.00e-24
 OTHER INFORMATION: SWISSPROT HIT: P56645, EVALUE 7.00e-24
 US-09-864-761-38417

Query Match 79.5%; Score 35; DB 9; Length 49;
 Best Local Similarity 66.7%; Pred. No. 73;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 7
 :|||||
 Db 22 HHHPA 27

RESULT 14
 US-09-272-975-52
 ; Sequence 52, Application US/09272975
 ; Publication No. US20030027774A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hendrickson, Ronald C.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Houghton, Raymond L.
 ; TITLE OF INVENTION: TUBERCULOSIS ANTIGENS AND METHODS
 ; TITLE OF INVENTION: OF USE THEREFOR
 ; FILE REFERENCE: 210121.474
 ; CURRENT APPLICATION NUMBER: US/09/272,975
 ; CURRENT FILING DATE: 1999-03-18
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 52
 ; LENGTH: 51
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-09-272-975-52

Query Match 79.5%; Score 35; DB 11; Length 51;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6
 :|||||
 Db 5 HHHP 9

RESULT 15
 US-10-369-493-4227
 ; Sequence 4227, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 4227
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 ; US-10-369-493-4227

Query Match 79.5%; Score 35; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6

Thu Feb 12 09:07:31 2004

us 03 201 1070 11111111

Db |||
 24 HIHP 28

Search completed: February 11, 2004, 17:54:13
Job time : 24.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THIHPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	88.6	312	2 D86488	bacterial IS-eleme
2	39	88.6	402	2 I67760	transposase - Esch
3	39	88.6	402	2 AE0737	probable transposa
4	37	84.1	66	2 C82702	hypothetical prote
5	37	84.1	69	2 C81075	hypothetical prote
6	37	84.1	611	2 T22855	hypothetical prote
7	37	84.1	722	2 S57246	ventral nervous sy
8	36	81.8	156	2 G81900	hypothetical prote
9	36	81.8	376	2 S45107	hypothetical prote
10	36	81.8	443	2 B81122	phage virion prote
11	36	81.8	827	2 T39608	zinc finger transc
12	35	79.5	169	2 A82736	hypothetical prote
13	35	79.5	250	2 C83397	pyrroloquinoline q
14	35	79.5	251	2 S20455	pqqC protein - Kle
15	35	79.5	617	2 A56051	myocyte nuclear fa
16	35	79.5	693	1 TNBEF7	73.8K alpha trans-
17	35	79.5	1234	2 A34911	band 3-related pro
18	35	79.5	1237	2 A31789	band 3-related pro
19	35	79.5	1237	2 A56764	band 3-related pro
20	35	79.5	1240	2 S21086	anion exchange pro
21	35	79.5	1530	2 S52239	brinfeldin a resist
22	35	79.5	1530	2 T52010	hypothetical prote
23	35	79.5	1820	2 A55494	latent transformin
24	34	77.3	108	2 AD1986	hypothetical prote
25	34	77.3	338	2 B64148	hypothetical prote
26	34	77.3	392	2 I39521	rubredoxin-NAD+ re
27	34	77.3	417	2 E83806	homoserine dehydro
28	34	77.3	434	2 AE3448	guanine deaminase
29	34	77.3	439	2 C65072	Guanine deaminase

30	34	77.3	439	2 D91098	hypothetical prote
31	34	77.3	439	2 H85943	hypothetical prote
32	34	77.3	474	2 S66480	carbon catabolite
33	34	77.3	515	2 H86202	hypothetical prote
34	34	77.3	520	2 T30808	hypothetical prote
35	34	77.3	566	2 A84485	hypothetical prote
36	34	77.3	704	2 S21911	BRcore-NS-Z3 prote
37	34	77.3	774	2 JC6095	hepatocyte nuclear
38	34	77.3	819	2 T10355	hypothetical prote
39	34	77.3	916	2 JT0396	reverse transcript
40	34	77.3	920	2 JC7313	aryl hydrocarbon r
41	34	77.3	1165	2 S27809	GTPase-activating
42	33	75.0	63	2 D82841	hypothetical prote
43	33	75.0	65	2 T18064	hypothetical prote
44	33	75.0	75	2 S25719	protein ECL - rat
45	33	75.0	250	2 T48684	hypothetical prote

ALIGNMENTS

RESULT 1

D86488
bacterial IS-element [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C;Accession: D86488
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141, MUID:21016719, PMID:11130712
A;Accession: D86488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <STO>
A;Cross-references: GB:AE005172; NID:G11094742; PIDN:AAG29676.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: Escherichia coli insertion sequence IS10 transposase

Query Match 88.6%; Score 39; DB 2; Length 312;
Best Local Similarity 85.7%; Pred. No. 7.8;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 157 THCHPS 163

RESULT 2

I67760
transposase - Escherichia coli insertion sequence IS10
C;Species: Escherichia coli
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 04-Mar-2000
C;Accession: I67760
R;Bogosian, G.; Billyeu, K.; O'Neil, J.P.
Gene 133, 17-22, 1993
A;Title: Genome rearrangements by residual IS10 elements in strains of Escherichia coli
A;Reference number: IS3672; MUID:94040791; PMID:8224890
A;Accession: I67760
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-402 <RES>
A;Cross-references: GB:S67119; NID:G455674; PIDN:AAB28848.1; PID:G455676
A;Experimental source: plasmid pX107; insertion sequence IS10
C;Genetics:

A;Mobile element: insertion sequence IS10
C;Superfamily: Escherichia coli insertion sequence IS10 transposase
C;Keywords: DNA binding

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THHHPS 7
|||
Db 247 THCHPS 253

RESULT 3

AF0737
Probable transposase STY2056 [imported] - Salmonella enterica serovar Typhimurium
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AF0737
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0737
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05600.1; PID:G16503097; GSPDB:GN00176
C;Genetics:
A;Gene: STY2056
C;Superfamily: Escherichia coli insertion sequence IS10 transposase

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THHHPS 7
|||
Db 247 THCHPS 253

RESULT 4

C82702
Hypothetical protein XF1277 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82702
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82702
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-66 <SIM>
A;Cross-references: GB:AE003961; GB:AE003849; NID:G9106254; PIDN:AAF84086.1; GSPDB:GN001
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1277

Query Match 84.1%; Score 37; DB 2; Length 66;
Best Local Similarity 85.7%; Pred. No. 3.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THHHPS 7
|||
Db 41 THDHPS 47

RESULT 5

C81075
Hypothetical protein NMB1503 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: C81075
R;Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: C81075
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-69 <TET>
A;Cross-references: GB:AE002500; GB:AE002098; NID:G7226744; PIDN:AAF41859.1; PID:G72267
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1503

Query Match 84.1%; Score 37; DB 2; Length 69;
Best Local Similarity 83.3%; Pred. No. 3.3;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 THHHPS 7
|||
Db 40 THHHPS 45

RESULT 6

T22855
Hypothetical protein F57F5.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22855
R;Harris, B.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19627
A;Accession: T22855
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-611 <WIL>
A;Cross-references: EMBL:Z75953; PIDN:CAB00100.1; GSPDB:GN00023; CESP:F57F5.4
A;Experimental source: clone F57F5
C;Genetics:
A;Gene: CESP:F57F5.4
A;Map position: 5
A;Introns: 26/3; 56/2; 101/1; 150/3; 296/3; 382/3; 486/3; 570/1

Query Match 84.1%; Score 37; DB 2; Length 611;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 THHHPS 7
|||
Db 451 THHHPS 456

RESULT 7

S57246
ventral nervous system defective protein - fruit fly (*Drosophila melanogaster*)
N/Alternate names: NK-2 homeotic protein
C/Species: *Drosophila melanogaster*
C/Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 15-Oct-1999
C/Accession: S57246; B33976
R/Jimenez, F.; Martin-Morris, L.E.; Velasco, L.; Chu, H.; Sierra, J.; Rosen, D.R.; White
EMBO J. 14, 3487-3495, 1995
A/Title: vnd, a gene required for early neurogenesis of *Drosophila*, encodes a homeodomain
A/Reference number: S57246; MUID:95354667; PMID:7628450
A/Accession: S57246
A/Molecule type: mRNA
A/Residues: 1-722 <JIM>
A/Cross-references: EMBL:X87141; NID:G1045047; PIDN:CAA60619.1; PID:G1045048
R/Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 86, 7716-7720, 1989
A/Title: *Drosophila* NK-homeobox genes.
A/Reference number: A33976; MUID:90046666; PMID:2573058
A/Accession: B33976
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 476-631, 'VG' <KIM>
A/Cross-references: GB:M27290; NID:G157635; PIDN:AAA28617.1; PID:G157636
C/Genetics:

A/Gene: FlyBase:vnd
A/Cross-references: FlyBase:FBgn0003986

C/Superfamily: unassigned homeobox proteins; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:545-601/Domain: homeobox homology <HOX>

Query Match 84.1%; Score 37; DB 2; Length 722;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 620 THPHPS 626

RESULT 8

G81900
hypothetical protein NMA1319 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup
C/Species: *Neisseria meningitidis*
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C/Accession: G81900
R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A/Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A/Reference number: A81775; MUID:20222556; PMID:10761919
A/Accession: G81900
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-156 <PAR>
A/Cross-references: GB:AL162755; GB:AL157959; NID:G7379742; PIDN:CAB84570.1; PID:G737999
A/Experimental source: serogroup A, strain Z2491
C/Genetics:

A/Gene: NMA1318; NMA1319

Query Match 81.8%; Score 36; DB 2; Length 156;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHPS 7

Db 138 HIHPA 143

RESULT 9

S45107

hypothetical protein 1 - *Erwinia carotovora*
C/Species: *Erwinia carotovora*
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C/Accession: S45107
R/Golby, P.; Jones, S.E.; Stephens, S.; Reeves, P.J.; Bycroft, B.; Stewart, G.; William
submitted to the EMBL Data Library, May 1994
A/Description: Global regulation of *Erwinia carotovora* exoenzyme virulence factors: mul
A/Reference number: S45107
A/Accession: S45107
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-376 <GOL>
A/Cross-references: EMBL:X79474; NID:G496597; PIDN:CAA55982.1; PID:G496598

Query Match 81.8%; Score 36; DB 2; Length 376;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHHP 6

Db 283 SHIHP 288

RESULT 10

B81122
phage virion protein, probable NMB1109 [imported] - *Neisseria meningitidis* (strain MCS8
C/Species: *Neisseria meningitidis*
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: B81122
R/Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A/Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MCS8.
A/Reference number: A81000; MUID:20175755; PMID:10710307
A/Accession: B81122
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-443 <TET>
A/Cross-references: GB:AE002460; GB:AE002098; NID:G7226335; PIDN:AAF41500.1; PID:G72263
A/Experimental source: serogroup B, strain MCS8
C/Genetics:

A/Gene: NMB1109

Query Match 81.8%; Score 36; DB 2; Length 443;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHPS 7

Db 425 HIHPA 430

RESULT 11

T39608
zinc finger transcription factor - fission yeast (*Schizosaccharomyces pombe*)
C/Species: *Schizosaccharomyces pombe*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C/Accession: T39608
R/Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
submitted to the EMBL Data Library, May 1998
A/Reference number: Z21866
A/Accession: T39608
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-827 <LYN>

A/Cross-references: EMBL:AL023554; PIDN:CAAL9035.1; GSPDB:GN00067; SPDB:SPBC16G5.16

A/Experimental source: strain 972h-; cosmid c16G5

C/Genetics:

A/Gene: SPDB:SPBC16G5.16

A/Map position: 2

A/Introns: 30/3; 41/1

C;Superfamily: GAL4 zinc binuclear cluster homology
F;11-47/Domain: GAL4 zinc binuclear cluster homology <GL4>

Query Match 81.8%; Score 36; DB 2; Length 827;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHPS 7
|||||
Db 221 HHHPA 226

RESULT 12

A82736
hypothetical protein XF1010 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: A82736
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82736
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <SIM>
A;Cross-references: GB:AE003938; GB:AE003849; NID:g9105935; PIDN:AAF83820.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1010

Query Match 79.5%; Score 35; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHP 6
|||||
Db 158 HHHHP 162

RESULT 13

C83397
pyrroloquinoline quinone biosynthesis protein C PA1987 [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: C83397
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, ; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: C83397
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-250 <STO>
A;Cross-references: GB:AE004625; GB:AE004091; NID:g9947983; PIDN:AAG05375.1; GSPDB:GN001 A;Experimental source: strain PA01

C;Genetics:
A;Gene: pqqC; PA1987

Query Match 79.5%; Score 35; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHP 6
|||||
Db 23 HHHHP 27

RESULT 14

S20455
pqqC protein - Klebsiella pneumoniae
C;Species: Klebsiella pneumoniae
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 08-Oct-1999
C;Accession: S20455; S21840
R;Meulenberg, J.J.M.; Sellink, E.; Riegman, N.H.; Postma, P.W.
Mol. Gen. Genet. 232, 284-294, 1992
A;Title: Nucleotide sequence and structure of the Klebsiella pneumoniae pqq operon.
A;Reference number: S20452; MUID:92212293; PMID:1313537
A;Accession: S20455
A;Molecule type: DNA
A;Residues: 1-251 <MEU>
A;Cross-references: EMBL:X58778; NID:g43903; PIDN:CAA41581.1; PID:g43907
C;Genetics:
A;Gene: pqqC

Query Match 79.5%; Score 35; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHP 6
|||||
Db 24 HHHHP 28

RESULT 15

A56051
myocyte nuclear factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Sep-1999
C;Accession: A56051
R;Bassel-Duby, R.; Hernandez, M.D.; Yang, Q.; Rochelle, J.M.; Seldin, M.F.; Williams, R Mol. Cell. Biol. 14, 4596-4605, 1994
A;Title: Myocyte nuclear factor, a novel winged-helix transcription factor under both d A;Reference number: A56051; MUID:94277065; PMID:8007964
A;Accession: A56051
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-617 <BAS>
A;Cross-references: GB:I26507; NID:g508527; PIDN:AAA37529.1; PID:g508528
C;Genetics:
A;Gene: MNF

C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
C;Keywords: DNA binding; transcription factor
F;289-380/Domain: fork head DNA-binding domain homology <FHD>

Query Match 79.5%; Score 35; DB 2; Length 617;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 THHHPS 7
|||||
Db 583 THSHPT 589

Search completed: February 11, 2004, 17:11:54
Job time : 11.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THIHPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	86.4	1878	1 BAZA_HUMAN	Q9uif9 homo sapien
2	37	84.1	341	1 BRN3_CHICK	Q91998 gallus gall
3	37	84.1	611	1 ADD2_CAEEL	Q20952 caenorhabdi
4	37	84.1	722	1 HNN2_DROME	P22808 drosophila
5	36	81.8	376	1 YQCB_ERWCA	Q47417 erwinia car
6	36	81.8	384	1 VPAP_PPVKA	P36702 pseudorabie
7	36	81.8	637	1 SCAG_CHICK	Q92075 gallus gall
8	35	79.5	251	1 PQOC_KLEPN	P27505 klebsiella
9	35	79.5	617	1 FXKI_MOUSE	P42128 mus musculu
10	35	79.5	693	1 UL47_HSV11	P10231 herpes simp
11	35	79.5	1210	1 PER3_HUMAN	P56645 homo sapien
12	35	79.5	1234	1 B3A2_RAT	P23347 rattus norv
13	35	79.5	1237	1 B3A2_MOUSE	P13808 mus musculu
14	35	79.5	1237	1 B3A2_RABIT	P48746 oryctolagus
15	35	79.5	1238	1 B3A2_CAVPO	Q92058 cavia porce
16	35	79.5	1241	1 B3A2_HUMAN	P04920 homo sapien
17	35	79.5	1530	1 BFR1_SCHPO	P41820 schizosacch
18	35	79.5	1850	1 BAZA_MOUSE	Q91ye5 mus musculu
19	34	77.3	338	1 YJEK_HAEIN	P44641 haemophilus
20	34	77.3	344	1 RURE_ACICA	Q96gd4 homo sapien
21	34	77.3	392	1 RURE_ACICA	P42454 acinetobact
22	34	77.3	408	1 HN4G_HUMAN	Q14541 homo sapien
23	34	77.3	408	1 HN4G_MOUSE	Q9wu06 mus musculu
24	34	77.3	439	1 GUAD_ECOLI	P76641 escherichia
25	34	77.3	474	1 MIG1_KLULA	P50898 kluyveromyc
26	34	77.3	819	1 Y083_NPVOP	O10336 orgyia pseu
27	34	77.3	916	1 RTJK_DROME	P21328 drosophila
28	34	77.3	1165	1 GAPI_DROME	P48423 drosophila
29	33	75.0	264	1 OTP_DROME	P56672 drosophila
30	33	75.0	297	1 CLPI_MOUSE	Q08091 mus musculu
31	33	75.0	297	1 CLPI_RAT	Q08290 rattus norv
32	33	75.0	337	1 Y020_BUCAP	Q44634 buchnera ap
33	33	75.0	439	1 GUAD_DEIRA	Q9ryx4 deinococcus

34 33 75.0 465 1 NQRA_CHLMU Q9plu3 chlamydia m
35 33 75.0 465 1 NQRA_CHLTR Q84639 chlamydia t
36 33 75.0 514 1 VNN1_CANFA Q9tsx8 canis famil
37 33 75.0 860 1 ATG6_HUMAN Q9ukp5 homo sapien
38 33 75.0 1275 1 RGA2_SCHPO Q10164 schizosacch
39 32 72.7 227 1 IFEL_ORYSA P48599 oryza sativ
40 32 72.7 368 1 VE2_HPV45 P36794 human papil
41 32 72.7 376 1 HMEV_DROME P06602 drosophila
42 32 72.7 449 1 GAT6_HUMAN Q92908 homo sapien
43 32 72.7 486 1 BAF1_KLULA P26375 kluyveromyc
44 32 72.7 518 1 VL2_HPV5B P26540 human papil
45 32 72.7 592 1 THDI_ARATH Q92896 arabidopsis

ALIGNMENTS

RESULT 1
BAZA_HUMAN
ID BAZA_HUMAN STANDARD; PRT; 1878 AA.
AC Q9UIF9; O00536; O15030; Q96H26;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bromodomain adjacent to zinc finger domain 2A (Transcription
DE termination factor-I interacting protein 5) (TTF-I interacting protein
DE 5) (Tips) (hwalp3).
GN BAZA2A OR TIP5 OR KIAA0314.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=20130112; PubMed=10662543;
RA Jones M.H., Hamana N., Nezu J.-I., Shimane M.;
RT "A novel family of bromodomain genes.";
RL Genomics 63:40-45(2000).
RN [2]
RP SEQUENCE OF 332-738 FROM N.A.
RC TISSUE=Lung;
RA Jansa P., Grummt I.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 639-1878 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [4]
RP SEQUENCE OF 1035-1878 FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Sosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

This SWISS-PROT entry is copyright. It is produced through a collaboration

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91998; CAA63049.1; --
 DR EMBL; X91997; CAA63048.1; --
 DR HSSP; P14859; 1OCT.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000327; POU domain.
 DR InterPro; IPR007103; POU_homeo.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00157; pou; 1.
 DR PRINTS; PR00028; POU; 1.
 DR ProDom; PD000010; Homeobox; 1.
 DR ProDom; PD000583; POU domain; 1.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00352; POU; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00035; POU_1; 1.
 DR PROSITE; PS00465; POU_2; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Transcription regulation; Nuclear protein.
 FT DOMAIN 55 65 POU-IV BOX.
 FT DOMAIN 185 259 POU.
 FT DNA BIND 277 336 HOMEBOX.
 SQ SEQUENCE 341 AA; 37773 MW; 6A4C741C88C07D5F CRC64;
 Query Match 84.1%; Score 37; DB 1; Length 341;
 Best Local Similarity 83.3%; Pred. No. 17;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHPS 7
 Db 95 HLHPS 100
 RESULT 3
 ADD2_CAEEL STANDARD; PRT; 611 AA.
 AC Q20952;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adducin-related protein 2.
 GN ADD-2 OR F57F5.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moorthy S., Bennett V.;
 RT "Molecular and functional analysis of the spectrin based membrane
 RT skeleton in Caenorhabditis elegans.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Harris B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: MEMBRANE-CYTOSKELETON-ASSOCIATED PROTEIN THAT PROMOTES
 CC THE ASSEMBLY OF THE SPECTRIN-ACTIN NETWORK (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO THE ALDOLASE CLASS II FAMILY. ADDUCIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF166171; AAD49860.1; --
 DR EMBL; Z75953; CAB00100.1; --
 DR PIR; T22855; T22855.
 DR WormPep; F57F5.4; CE06002.
 DR InterPro; IPR001303; Aldolase II_N.
 DR Pfam; PF00596; Aldolase_II; 1.
 KW Cytoskeleton; Membrane.
 SQ SEQUENCE 611 AA; 68561 MW; 9815707C45CD70D8 CRC64;
 Query Match 84.1%; Score 37; DB 1; Length 611;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HHHPS 7
 Db 451 HLHPS 456
 RESULT 4
 HMN2_DROME STANDARD; PRT; 722 AA.
 AC P22808; Q24589;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Homeobox protein vnd (ventral nervous system defective protein)
 DE (Homeobox protein NK-2).
 GN VND OR NK2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95354667; PubMed=7628450;
 RA Jimenez F., Martin-Morris L.E., Velasco L., Chu H., Sierra J.,
 RA Rosen D.R., White K.;
 RT "vnd, a gene required for early neurogenesis of Drosophila, encodes a
 RT homeodomain protein.";
 RL EMBO J. 14:3487-3495(1995).
 RN [2]
 RP SEQUENCE OF 476-633 FROM N.A.
 RX MEDLINE=90046666; PubMed=2573058;
 RA Kim Y., Nirenberg M.;
 RT "Drosophila NK-homeobox genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7716-7720(1989).
 RN [3]
 RP STRUCTURE BY NMR OF 537-612.
 RX MEDLINE=95371127; PubMed=7643404;
 RA Tsao D.H.H., Gruschus J.M., Wang L.-H., Nirenberg M., Ferretti J.A.;
 RT "The three-dimensional solution structure of the NK-2 homeodomain
 RT from Drosophila.";
 RL J. Mol. Biol. 251:297-307(1995).
 RN [4]
 RP STRUCTURE BY NMR OF 543-605.
 RX MEDLINE=97299770; PubMed=9154919;
 RA Gruschus J.M., Tsao D.H.H., Wang L.-H., Nirenberg M., Ferretti J.A.;
 RT "Interactions of the vnd/NK-2 homeodomain with DNA by nuclear magnetic
 RT resonance spectroscopy: basis of binding specificity.";
 RL Biochemistry 36:5372-5380(1997).
 CC -|- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN THE
 CC REGULATION OF THE PRONEURAL AS-C GENES AND THE NEUROGENIC GENES OF
 CC THE ENHANCER OF SPLIT COMPLEX. COULD SPECIFICALLY ACTIVATE
 CC PRONEURAL GENES IN THE VENTRAL-MOST NEUROECTODERM.
 CC -|- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -|- TISSUE SPECIFICITY: EXPRESSED IN THE CNS AND MIDGUT.
 CC -|- SIMILARITY: BELONGS TO THE NK-2 HOMEBOX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X87141; CAA60619.1; -
DR EMBL; M27290; AAA28617.1; -
DR PIR; S57246; S57246.
DR PDB; 1VND; 08-NOV-96.
DR PDB; 1NK2; 23-FEB-99.
DR PDB; 1NK3; 23-FEB-99.
DR PDB; 1QRY; 06-JUL-99.
DR TRANSFAC; T04258; -
DR FlyBase; FBgn0003986; vnd.
DR GO; GO:0007400; P:neuroblast cell fate determination; IMP.
DR GO; GO:0007419; P:ventral cord development; NAS.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; Hox; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.

KW Transcription regulation; Homeobox; DNA-binding; 3D-structure;
KW Developmental protein; Nuclear protein; Neurogenesis.
FT DOMAIN 108 111 POLY-SER.
FT DOMAIN 215 223 POLY-ALA.
FT DOMAIN 336 343 POLY-ALA.
FT DOMAIN 351 359 POLY-ALA.
FT DOMAIN 482 486 POLY-ASN.
FT DOMAIN 489 492 POLY-ASN.
FT DOMAIN 544 603 HOMEBOX.
FT DOMAIN 688 693 POLY-ALA.
FT VARIANT 578 578 A -> T (IN VND29).
FT CONFLICT 632 633 RR -> VG (IN REF. 2).
FT TURN 553 554
FT HELIX 555 565
FT HELIX 571 581
FT TURN 582 582
FT HELIX 585 594

SQ SEQUENCE 722 AA; 76468 MW; D036AE4D890014DA CRC64;
Query Match 84.1%; Score 37; DB 1; Length 722;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 THIHHP 7
DB 620 THPHHP 626

RESULT 5
YQCB_ERWCA
ID YQCB_ERWCA STANDARD; PRT; 376 AA.
AC Q47417;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exoenzyme regulation regulon ORF1.
OS Erwinia carotovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI193;
RA Golby P., Jones S.E., Stephens S., Reeves P.J., Bycroft B.,
RA Stewart G., Williams P., Salmon G.P.C.;
RT "Global regulation of Erwinia carotovora exoenzyme virulence
RT factors: multicopy suppression of rex mutants and evidence for a
RT global repression regulon.";
RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE RLJ FAMILY OF PSEUDOURIDINE SYNTHASES.
CC STRONG, TO E.COLI YQCB.
CC -!- SIMILARITY: IN THE N-TERMINAL, STRONG, TO E.COLI YQCC; IN THE C-
CC TERMINAL, STRONG, TO E.COLI YQCB.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X79474; CAA55982.1; -
DR PIR; S45107; S45107.
DR InterPro; IPR006145; PseudoU synth.
DR InterPro; IPR006224; PSI_RLJ.
DR Pfam; PF04287; DUF446; 1.
DR Pfam; PF00849; PseudoU synth_2; 1.
DR ProDom; PD001819; PSI_RLJ; 1.
DR PROSITE; PS01129; PSI_RLJ; 1.
SQ SEQUENCE 376 AA; 43605 MW; F35992CAAAD22E30 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 376;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 THIHHP 6
DB 283 SHIHHP 288

RESULT 6
VPAP_PVKKA
ID VPAP_PVKKA STANDARD; PRT; 384 AA.
AC P36702;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA polymerase processivity factor (Polymerase accessory protein)
DE (PAP) (UL42 homolog).
OS Pseudorabies virus (strain Kaplan) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=33703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95222727; PubMed=7707503;
RA Berthomme H., Monahan S.J., Parris D.S., Jacquemont B.,
RA Epstein A.L.;
RT "Cloning, sequencing, and functional characterization of the two
RT subunits of the pseudorabies virus DNA polymerase holoenzyme:
RT evidence for specificity of interaction.";
RL J. Virol. 69:2811-2818 (1995).
CC -!- FUNCTION: ACCESSORY SUBUNIT OF THE DNA POLYMERASE THAT ACTS TO
CC INCREASE THE PROCESSIVITY OF POLYMERIZATION (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES POLYMERASE ACCESSORY PROTEIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M94355; AAA74384.1; -
DR InterPro; IPR003202; UL42.
DR Pfam; PF02282; UL42; 2.
KW DNA-binding; DNA replication.
SQ SEQUENCE 384 AA; 40305 MW; FFAC3EF0C1984936 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 384;

```
Best Local Similarity 71.4%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 16 THAHHPA 22

RESULT 7
SCAA_CHICK STANDARD; PRT; 637 AA.
AC Q92075; P70095; Q98941;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Amiloride-sensitive sodium channel alpha-subunit (Epithelial Na+
channel alpha subunit) (Alpha ENaC) (Nonvoltage-gated sodium channel
1 alpha subunit) (SCNEA) (Alpha NaCH).
GN SCNN1A OR ENAC.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC STRAIN=Isa brown; TISSUE=Cochlea;
RX MEDLINE=97157073; PubMed=9003454;
RA Killick R., Richardson G.;
RT "Isolation of chicken alpha ENaC splice variants from a cochlear cDNA
library.";
RL Biochim. Biophys. Acta 1350:33-37(1997).
[2]
SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Goldstein O., Asher C., Garty H.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SODIUM PERMEABLE NON-VOLTAGE-SENSITIVE ION CHANNEL
INHIBITED BY THE DIURETIC AMILORIDE. MEDIATE THE ELECTRODIFFUSION
OF THE LUMINAL SODIUM (AND WATER, WHICH FOLLOWS OSMOTICALLY)
THROUGH THE APICAL MEMBRANE OF EPITHELIAL CELLS. CONTROLS THE
REABSORPTION OF SODIUM IN KIDNEY, COLON, LUNG AND SWEAT GLANDS.
CC ALSO PLAYS A ROLE IN TASTE PERCEPTION (BY SIMILARITY).
CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA, ONE BETA AND ONE GAMMA
SUBUNIT. A DELTA SUBUNIT CAN REPLACE THE ALPHA SUBUNIT (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Long;
Isoid=Q92075-1; Sequence=Displayed;
Name=Short;
Isoid=Q92075-2; Sequence=VSP 006196, VSP 006197;
CC TISSUE SPECIFICITY: THE LONG ISOFORM HAS BEEN FOUND IN COCHLEA,
COLON, AND CARTILAGE. THE SHORT ISOFORM IS ONLY FOUND IN COCHLEA.
CC -!- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNEL
FAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-15 IS THE INITIATOR.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U62902; AAB50550.1; --
EMBL; U62903; AAB50551.1; --
EMBL; U62904; AAB50552.1; --
EMBL; U58475; AAB04954.1; ALT_INIT.
DR InterPro; IPR004724; EnaC.
DR InterPro; IPR001873; Na+channel_ASC.
```

```
* DR Pfam; PF00858; ASC; 1.
DR PRINTS; PRO1078; AMINACHANNEL.
DR TIGRFS; TIGR00859; ENaC; 1.
DR PROSITE; PS01206; ASC; 1.
KW Ionic channel; Transmembrane; Ion transport; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 80
FT TRANSMEM 81 97
FT DOMAIN 98 554
FT TRANSMEM 555 571
FT DOMAIN 572 637
FT CARBOHYD 157 157
FT CARBOHYD 280 280
FT CARBOHYD 298 298
FT CARBOHYD 499 499
FT CARBOHYD 402 434
FT VARSPLIC
FT FT
FT FT
FT VARSPLIC 435 637
FT CONFLICT 49 50
FT CONFLICT 578 578
FT SEQUENCE 637 AA; 73236 MW; 5D15E616373971DA CRC64;
SQ
Query Match 81.8%; Score 36; DB 1; Length 637;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHHP 6
Db 282 THFHHP 287

RESULT 8
PQQC_KLEPN STANDARD; PRT; 251 AA.
ID PQQC_KLEPN
AC P27505;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Coenzyme PQQ synthesis protein C.
GN PQQC.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 418;
RX MEDLINE=92212293; PubMed=1313537;
RA Meulenbergh J.J.M., Sellink E., Riegman N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae pqq
operon.";
RL Mol. Gen. Genet. 232:284-294(1992).
CC -!- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
BIOSYNTHESIS.
CC -!- SIMILARITY: TO OTHER BACTERIAL PQQC.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X58778; CAA41581.1; --
DR PIR; S20455; S20455.
DR Pfam; PF05312; PQQC; 1.
KW PQQ.
SQ SEQUENCE 251 AA; 28986 MW; B48A494FA63B6598 CRC64;
```


Query Match 79.5%; Score 35; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHP 6
Db 24 HIHP 28

RESULT 9
FXK1_MOUSE STANDARD; PRT; 617 AA.
AC P42128; O35939;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Forkhead box protein K1 (Myocyte nuclear factor) (MNF).
GN FOXK1 OR MNF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=94277065; PubMed=8007964;
RA Bassel-Duby R., Hernandez M.D., Yang Q., Rochelle J.M.,
RA Seldin M.F., Williams R.S.;
RT "Myocyte nuclear factor, a novel winged-helix transcription factor
under both developmental and neural regulation in striated
myocytes.";
RL Mol. Cell. Biol. 14:4596-4605(1994).
RN [2]
SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=97415602; PubMed=9271401;
RA Yang Q., Bassel-Duby R., Williams R.S.;
RT "Transient expression of a winged-helix protein, MNF-beta, during
myogenesis.";
RL Mol. Cell. Biol. 17:5236-5243(1997).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO THE UPSTREAM
ENHANCER REGION (CCAC BOX) OF MYOGLOBIN GENE. HAS A ROLE IN
MYOGENIC DIFFERENTIATION AND IN REMODELING PROCESSES OF ADULT
MUSCLES THAT OCCUR IN RESPONSE TO PHYSIOLOGICAL STIMULI.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P42128-1; Sequence=Displayed;
CC Name=2; Synonyms=Beta;
CC IsoId=P42128-2; Sequence=VSP 001545, VSP 001546;
CC -!- TISSUE SPECIFICITY: EXPRESSED IN TISSUES AND CELLS IN WHICH THE
MYOGLOBIN GENE IS TRANSCRIPTIONALLY ACTIVE (CARDIAC AND SKELETAL
MYOCYTES, BRAIN, KIDNEY. . .).
CC -!- PTM: PHOSPHORYLATED.
CC -!- SIMILARITY: Contains 1 fork-head domain.
CC -!- SIMILARITY: Contains 1 FHA domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; L26507; AAA37529.1; -
EMBL; U95016; AAB69641.1; -
DR PIR; A56051; A56051.
DR HSSP; Q63245; 2HFH.
DR TRANSFAC; T04216; -
DR MGD; MGI:1347488; Foxk1.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00498; FHA; 1.

DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PRO0053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW phosphorylation; Alternative splicing; Differentiation.
FT DOMAIN 27 33 POLY-ALA.
FT DOMAIN 107 159 FHA.
FT DNA_BIND 288 379 FORK-HEAD.
FT VARSPLIC 400 409 RSPASPTHPG -> SAPASHTSHA (in isoform 2).
FT VARSPLIC 410 617 /FTId=VSP_001545.
FT VARSPLIC 410 617 Missing (in isoform 2).
FT VARSPLIC 410 617 /FTId=VSP_001546.
SQ SEQUENCE 617 AA; 65839 MW; A1083B28C709FC4A CRC64;
Query Match 79.5%; Score 35; DB 1; Length 617;
Best Local Similarity 71.4%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 583 THSHPT 589

RESULT 10
UL47_HSV11 STANDARD; PRT; 693 AA.
ID UL47_HSV11
AC P10231;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Virion protein UL47 (82/81 kDa tegument protein) (VMW82/81) (VPI3/14).
GN UL47.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
herpes simplex virus type 1.";
RL J. Gen. Virol. 69:1531-1574(1988).
RN [2]
IDENTIFICATION OF PROTEIN.
RX MEDLINE=91108391; PubMed=2177087;
RA McLean G., Rixon F., Langeland N., Haarr L., Marsden H.;
RT "Identification and characterization of the virion protein products
of herpes simplex virus type 1 gene UL47.";
RL J. Gen. Virol. 71:2953-2960(1990).
CC -!- FUNCTION: MODULATOR OF ALPHA-TIF (VMW65 PHOSPHOPROTEIN) TRANS-
ACTIVATION. POSSIBLY UL47 MAY HAVE KINASE ACTIVITY.
CC -!- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
EHV-1 13, AND VZV 11.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X14112; CAA32297.1; -
DR


```

CC -1- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
CC DISTRIBUTION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05166; AAA40799.1; -.
CC PIR; A34911; A34911.
CC HSSP; P02730; 1BTQ.
CC InterPro; IPR001717; Anion_exchange.
CC InterPro; IPR003020; HCO3_cotranspt.
CC Pfam; PF00955; HCO3_cotransp; 1.
CC PRINTS; PR01231; HCO3TRNSPORT.
CC TIGRFAMs; TIGR00834; ae; 1.
CC PROSITE; PS00219; ANION_EXCHANGER_1; 1.
CC PROSITE; PS00220; ANION_EXCHANGER_2; 1.
CC Transmembrane; Glycoprotein; Transport; Antiport; Ion transport;
CC Anion exchange; Lipoprotein; Palmitate.
CC DOMAIN 1 704 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 2 705 MEMBRANE (ANION EXCHANGE).
CC TRANSMEM 705 728 POTENTIAL.
CC TRANSMEM 734 771 POTENTIAL.
CC TRANSMEM 791 813 POTENTIAL.
CC TRANSMEM 823 844 POTENTIAL.
CC TRANSMEM 844 893 EXOPLASMIC LOOP (POTENTIAL).
CC TRANSMEM 894 911 POTENTIAL.
CC TRANSMEM 912 926 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 927 947 POTENTIAL.
CC TRANSMEM 981 1003 POTENTIAL.
CC TRANSMEM 1029 1050 POTENTIAL.
CC TRANSMEM 1084 1129 POTENTIAL.
CC TRANSMEM 1156 1192 POTENTIAL.
CC TRANSMEM 1192 1192 PRO-RICH.
CC DOMAIN 5 317 HIS-RICH.
CC DOMAIN 74 88 N-LINKED (GLNAC. . .) (POTENTIAL).
CC CARBOHYD 856 856 N-LINKED (GLNAC. . .) (POTENTIAL).
CC CARBOHYD 866 866 N-LINKED (GLNAC. . .) (POTENTIAL).
CC CARBOHYD 878 878 N-LINKED (GLNAC. . .) (POTENTIAL).
CC LIPID 1166 1166 PALMITATE (BY SIMILARITY).
CC CONFLICT 206 206 G -> A (IN REF. 2).
CC CONFLICT 925 926 RR -> PG (IN REF. 2).
CC CONFLICT 1018 1018 M -> I (IN REF. 2).
CC CONFLICT 1156 1156 M -> I (IN REF. 2).
CC SEQUENCE 1234 AA; 136635 MW; FAB4ED12BB916216 CRC64;
SQ
Query Match 79.5%; Score 35; DB 1; Length 1234;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HIHP 6
Db 80 HIHP 84
RESULT 13
B3A2_MOUSE
ID B3A2_MOUSE STANDARD; PRT; 1237 AA.
AC P13808; Q9ES09; Q9ES10; Q9ES11; Q9ES12; Q9ES13;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anion exchange protein 2 (Non-erythroid band 3-like protein) (B3BP).
GN SLC4A2 OR AE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=89034212; PubMed=3182834;
RA Alper S.L., Kopito R.R., Libresco S.M., Lodish H.F.;
RT "Cloning and characterization of a murine band 3-related cDNA from
RT kidney and from a lymphoid cell line.";
RL J. Biol. Chem. 263:17092-17099(1988).
RN [2]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=20462926; PubMed=11006093;
RA Lecanda J., Urtasun R., Medina J.F.;
RT "Molecular cloning and genomic organization of the mouse AE2 anion
RT exchanger gene.";
RL Biochem. Biophys. Res. Commun. 276:117-124(2000).
CC -1- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
CC DISTRIBUTION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=A;
CC IsoId=P13808-1; Sequence=Displayed;
CC Name=B1;
CC IsoId=P13808-2; Sequence=VSP_000458;
CC Name=B2;
CC IsoId=P13808-3; Sequence=VSP_000457;
CC Name=C1;
CC IsoId=P13808-4; Sequence=VSP_000460;
CC Name=C2;
CC IsoId=P13808-5; Sequence=VSP_000459, VSP_000461;
CC -1- TISSUE SPECIFICITY: Isoform a is widely expressed at similar
CC levels in all tissues examined. Isoforms b1 and b2 are
CC predominantly expressed in stomach although they are also detected
CC at lower levels in other tissues. Isoform c1 is stomach-specific.
CC Isoform c2 is expressed at slightly higher levels in lung and
CC stomach than in other tissues.
CC -1- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04036; AAA65505.1; -.
CC EMBL; AF255774; AAG23154.1; -.
CC EMBL; AF255774; AAG23155.1; -.
CC EMBL; AF255774; AAG23156.1; -.
CC EMBL; AF255774; AAG23158.1; -.
CC EMBL; AF255774; AAG23157.1; -.
CC PIR; A31789; A31789.
CC HSSP; P02730; 1BTQ.
CC MGD; MGI:109351; Slc4a2.
CC InterPro; IPR001717; Anion_exchange.
CC InterPro; IPR003020; HCO3_cotranspt.
CC Pfam; PF00955; HCO3_cotransp; 1.
CC PRINTS; PR01231; HCO3TRNSPORT.
CC TIGRFAMs; TIGR00834; ae; 1.
CC PROSITE; PS00219; ANION_EXCHANGER_1; 1.
CC PROSITE; PS00220; ANION_EXCHANGER_2; 1.
CC Transmembrane; Glycoprotein; Transport; Antiport; Ion transport;
CC Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
CC DOMAIN 1 703 MEMBRANE (ANION EXCHANGE).
CC DOMAIN 2 704 1237 POTENTIAL.
CC TRANSMEM 704 727 POTENTIAL.
CC TRANSMEM 733 770 POTENTIAL.
CC TRANSMEM 790 812 POTENTIAL.
CC TRANSMEM 822 843 POTENTIAL.
CC TRANSMEM 844 896 EXOPLASMIC LOOP (POTENTIAL).
CC TRANSMEM 897 914 POTENTIAL.
CC TRANSMEM 915 929 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 930 950 POTENTIAL.

```

FT TRANSMEM 984 1006 POTENTIAL.
FT TRANSMEM 1032 1053 POTENTIAL.
FT TRANSMEM 1087 1132 POTENTIAL.
FT TRANSMEM 1159 1195 POTENTIAL.
FT DOMAIN 5 316 PRO-RICH.
FT DOMAIN 73 87 HIS-RICH.
FT DOMAIN 861 865 POLY-SER.
FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 878 878 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 1169 1169 PALMITATE (BY SIMILARITY).
FT VARSPLIC 1 17 MSSAPRRPASGADSLHT -> MDFLLRPQ (in isoform B2).
FT VARSPLIC 1 17 /FTid=VSP 000457.
FT VARSPLIC 1 17 MSSAPRRPASGADSLHT -> MTQ (in isoform B1).
FT VARSPLIC 1 166 /FTid=VSP 000458.
FT VARSPLIC 1 198 /FTid=VSP 000459.
FT VARSPLIC 1 193 Missing (in isoform C1).
FT VARSPLIC 167 193 /FTid=VSP 000460.
FT CONFLICT 205 205 ERTSPPTQTPHQEAPRASKGAQTG -> MPAFQEWKSG
FT SEQUENCE 1237 AA; 136813 MW; 1A0782C0071782EE CRC64; GLREAVFGAGCSVCR (in isoform C2).
FT /FTid=VSP 000461.
FT A -> G (IN REF. 2).
FT CONFLICT 205 205
FT SEQUENCE 1237 AA; 136813 MW; 1A0782C0071782EE CRC64;
Query Match 79.5%; Score 35; DB 1; Length 1237;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HIHHP 6
Db 79 HIHHP 83
RESULT 14
B3A2_RABIT
ID B3A2_RABIT STANDARD; PRT; 1237 AA.
AC P48746;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anion exchange protein 2 (Non-erythroid band 3-like protein) (B3RP).
GN SLC4A2 OR AE2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white;
RX MEDLINE=93035730; PubMed=1415547;
RA Chow A., Dobbins J.W., Aronson P.S., Igarashi P.;
RT "cdna cloning and localization of a band 3-related protein from ileum."
RL Am. J. Physiol. 263:G345-G352(1992).
CC -!- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE DISTRIBUTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S45791; AAB23488.1; -.
CC PIR; A56764; A56764.
CC HSSP; P02730; 1BTQ.
DR InterPro; IPR001717; Anion_exchange.

DR InterPro; IPR003020; HCO3_cotranspt.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR01231; HCO3TRNSPORT.
DR TIGRFAMS; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
KW Transmembrane; Glycoprotein; Transport; Antiport; Ion transport;
KW Anion exchange; Lipoprotein; Palmitate.
FT DOMAIN 1 703 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 704 1237 MEMBRANE (ANION EXCHANGE).
FT TRANSMEM 704 727 POTENTIAL.
FT TRANSMEM 733 770 POTENTIAL.
FT TRANSMEM 790 812 POTENTIAL.
FT TRANSMEM 822 843 POTENTIAL.
FT TRANSMEM 844 896 EXOPLASMIC LOOP (POTENTIAL).
FT TRANSMEM 897 914 POTENTIAL.
FT TRANSMEM 915 929 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 930 950 POTENTIAL.
FT TRANSMEM 984 1006 POTENTIAL.
FT TRANSMEM 1032 1053 POTENTIAL.
FT TRANSMEM 1087 1132 POTENTIAL.
FT TRANSMEM 1159 1195 POTENTIAL.
FT TRANSMEM 1159 1195 PRO-RICH.
FT DOMAIN 5 316 HIS-RICH.
FT DOMAIN 74 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 864 864 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 878 878 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 1169 1169 PALMITATE (BY SIMILARITY).
FT SEQUENCE 1237 AA; 136535 MW; 2811D11051552BB2 CRC64;
Query Match 79.5%; Score 35; DB 1; Length 1237;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HIHHP 6
Db 80 HIHHP 84
RESULT 15
B3A2_CAVPO
ID B3A2_CAVPO STANDARD; PRT; 1238 AA.
AC Q9Z0S8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anion exchange protein 2 (Non-erythroid band 3-like protein) (AE2 anion exchanger).
DE SLC4A2 OR AE2.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH 2; TISSUE=Organ of Corti;
RX MEDLINE=99023787; PubMed=9804866;
RA Mhatre A.N., Charachon G., Alper A.L., Lalwani A.K.;
RT "The guinea pig cochlear AE2 anion exchanger: cdna cloning and in situ localization within the cochlea."
RL Biochim. Biophys. Acta 1414:1-15(1998).
CC -!- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE DISTRIBUTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC

```
-----
CC EMBL; AF121253; AAD19700.1; -.
DR HSP; P02730; 1BTQ.
DR InterPro; IPR001717; Anion_exchange.
DR InterPro; IPR003020; HCO3_cotranspt.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PR01231; HCO3TRNSPORT.
DR TIGRFAMs; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
KW Transmembrane; Glycoprotein; Transport; Antiport; Ion transport;
KW Anion exchange; Lipoprotein; Palmitate.
FT DOMAIN 1 704 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 705 1238 MEMBRANE (ANION EXCHANGE).
FT TRANSMEM 705 728 POTENTIAL.
FT TRANSMEM 734 771 POTENTIAL.
FT TRANSMEM 791 813 POTENTIAL.
FT TRANSMEM 823 844 POTENTIAL.
FT DOMAIN 845 897 EXOPLASMIC LOOP (POTENTIAL).
FT TRANSMEM 898 915 POTENTIAL.
FT DOMAIN 916 930 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 931 951 POTENTIAL.
FT TRANSMEM 985 1007 POTENTIAL.
FT TRANSMEM 1033 1054 POTENTIAL.
FT TRANSMEM 1088 1133 POTENTIAL.
FT TRANSMEM 1160 1196 POTENTIAL.
FT DOMAIN 5 316 PRO-RICH.
FT DOMAIN 74 88 HIS-RICH.
FT CARBOHYD 856 856 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 865 865 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 879 879 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 1170 1170 PALMITATE (BY SIMILARITY).
SQ SEQUENCE 1238 AA; 137358 MW; FA1739862EDSADBF CRC64;

Query Match 79.5%; Score 35; DB 1; Length 1238;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHP 6
Db 80 HIHP 84
```

Search completed: February 11, 2004, 17:04:16
Job time : 7.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THIHPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	95.5	468	10	O81027 arabidopsis
2	40	90.9	1202	5	O81DT1 plasmodium
3	39	88.6	73	5	O8SY32 drosophila
4	39	88.6	312	10	O9C7M4 arabidopsis
5	39	88.6	402	2	O53371 escherichia
6	39	88.6	402	2	O8VSK3 shigella fl
7	39	88.6	402	2	O9K388 salmonella
8	39	88.6	402	2	O8RQK7 bacillus ce
9	39	88.6	402	2	O56368 escherichia
10	39	88.6	402	2	O8GC95 citrobacter
11	39	88.6	402	4	O9H5N6 homo sapien
12	39	88.6	402	4	O9BTF6 homo sapien
13	39	88.6	402	5	O95TW0 drosophila
14	39	88.6	402	5	O95T96 drosophila
15	39	88.6	402	12	O8QTC0 white spot
16	39	88.6	402	12	O9E8P4 bluetongue

17	39	88.6	402	16	Q9S459	Q9s459 salmonella
18	39	88.6	559	10	Q947U9	Q947u9 oryza sativ
19	38	86.4	308	5	O76762	O76762 anopheles g
20	38	86.4	393	5	O8WSF5	O8wsf5 plasmodium
21	38	86.4	393	5	O8I537	O8i537 plasmodium
22	38	86.4	470	10	O8LX5	O8lgx5 oryza sativ
23	38	86.4	1377	5	O9W2F3	O9w2f3 drosophila
24	37	84.1	66	16	Q9PDV2	Q9pdv2 xylella fas
25	37	84.1	69	16	Q9JYN1	Q9jyn1 neisseria m
26	37	84.1	112	4	O8N377	O8n377 homo sapien
27	37	84.1	239	13	O8JHA2	O8jha2 serinus can
28	37	84.1	260	2	Q93P79	Q93p79 microscilla
29	37	84.1	343	13	Q9DG50	Q9dg50 xenopus lae
30	37	84.1	700	3	O8X007	O8x007 neurospora
31	37	84.1	723	5	Q9UB43	Q9ub43 drosophila
32	37	84.1	723	5	Q9W5F0	Q9w5f0 drosophila
33	37	84.1	723	5	O26436	O26436 drosophila
34	37	84.1	840	11	O8K062	O8k062 mus musculu
35	36	81.8	156	16	Q9JUG8	Q9jug8 neisseria m
36	36	81.8	184	16	O8VKL1	O8vkl1 mycobacteri
37	36	81.8	443	16	O9JZC9	O9jzc9 neisseria m
38	36	81.8	553	10	O9ZR78	O9zr78 nicotiana s
39	36	81.8	827	3	O60130	O60130 schizosacch
40	36	81.8	869	5	Q9Y0Z8	Q9y0z8 drosophila
41	36	81.8	1041	5	O8WTI7	O8wti7 drosophila
42	36	81.8	1041	5	O9VHL6	O9vhl6 drosophila
43	35	79.5	169	16	Q9PEL8	Q9pel8 xylella fas
44	35	79.5	250	16	Q9I2C2	Q9i2c2 pseudomonas
45	35	79.5	271	16	O8PFA4	O8pfa4 xanthomonas

ALIGNMENTS

RESULT 1

O81027	ID	O81027	PRELIMINARY;	PRT;	468	AA.
AC	O81027; Q9PPE2;					
DT	01-NOV-1998 (TREMBlrel. 08, Created)					
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)					
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)					
DE	Putative hydroxymethylglutaryl-CoA lyase (Putative					
DE	hydroxymethylglutaryl-CoA lyase protein).					
GN	AT2G26800.					
OS	Arabidopsis thaliana (Mouse-ear cress).					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;					
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi					
OX	NCBI_TaxID=3702;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=cv. Columbia;					
RA	Rounsley S.D., Ronning C.M., Lin X., Ketchum K.A., Crosby M.L.,					
RA	Brandon R.C., Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R.,					
RA	Adams M.D., Somerville C.R., Venter J.C.;					
RL	Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=cv. Columbia;					
RA	Lin X.;					
RL	Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=cv. Columbia;					
RA	Town C.D., Kaul S.;					
RL	Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RA	Sakano H., Liu S.X., Pham P.K., Yamada K., Banh J., Etgu P., Lee J.M.,					
RA	Toriumi M., Yu G., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;					
RT	"Full Length cDNA of gene F12C20.16/At2g26800 (GI:3426048).";					
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.					
RN	[5]					

RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninici P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL "Full Length cDNA of gene F12C20.16/At2g26800 (GI:3426048).";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AC005168; AAC32247.2; -;
DR EMBL; AF327420; AAG42010.1; -;
DR EMBL; AF349521; AAK15568.1; -;
DR InterPro; IPR000891; HMGL-like.
DR Pfam; PF00682; HMGL-like; 1.
KW Lyase.
SQ SEQUENCE 468 AA; 50577 MW; C91C70DAD241AF74 CRC64;

Query Match 95.5%; Score 42; DB 10; Length 468;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 21 THLHPS 27

RESULT 2
Q8IDT1 PRELIMINARY; PRT; 1202 AA.
AC Q8IDT1
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF13_0221.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,
RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AL844509; CAD52538.1; -;
KW Hypothetical protein.
SQ SEQUENCE 1202 AA; 146276 MW; D89DAA2506A1071F CRC64;

Query Match 90.9%; Score 40; DB 5; Length 1202;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHHP 6
Db 294 THIHHP 299

RESULT 3
Q8SY32 PRELIMINARY; PRT; 73 AA.
AC Q8SY32
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE LP02734P.
GN BCDNA:LP02734.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY075431; AAL68246.1; -;
DR FlyBase; FBgn0061414; BCDNA:LP02734.
SQ SEQUENCE 73 AA; 8338 MW; B1862F88AA421AA3 CRC64;

Query Match 88.6%; Score 39; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHPS 7
Db 39 HIHPS 44

RESULT 4
Q9C7M4 PRELIMINARY; PRT; 312 AA.
AC Q9C7M4
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Bacterial IS-element.
GN F103.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Pederspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC068901; AAG50897.1; -;
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
SQ SEQUENCE 312 AA; 35554 MW; 3C06A8F768CE29E8 CRC64;

Query Match 88.6%; Score 39; DB 10; Length 312;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 157 THCHPS 163

```
RESULT 5
Q53371 ID Q53371 PRELIMINARY; PRT; 402 AA.
AC Q53371;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Transposase (IS10 transposase).
GN R0085.
OS Escherichia coli, and
OS Salmonella typhi.
OG Plasmid pXT107, and Plasmid R27.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 601;
RN [1]
SEQUENCE FROM N.A.
RC SPECIES=E.coli; PLASMID=pXT107;
RX MEDLINE=94040791; PubMed=8224890;
RA Bogosian G., Bilyeu K., O'Neil J.P.;
RT "Genome rearrangements by residual IS10 elements in strains of
RT Escherichia coli K-12 which had undergone Tn10 mutagenesis and fusaric
RT acid selection.";
RL Gene 133:17-22(1993).
RN [2]
SEQUENCE FROM N.A.
RC SPECIES=S.typhi; PLASMID=R27;
RX MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grotbeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
RN [3]
SEQUENCE FROM N.A.
RC SPECIES=S.typhi; PLASMID=R27;
RX MEDLINE=20247426; PubMed=10783303;
RA Lawley T.D., Burland V.D., Taylor D.E.;
RT "Analysis of the complete nucleotide sequence of the tetracycline-
RT resistance transposon Tn10.";
RL Plasmid 43:235-239(2000).
DR EMBL; S67119; AAB28848.1; -.
DR EMBL; AF250878; AAF69923.1; -.
DR EMBL; AF223162; AAF68939.1; -.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Plasmid.
SQ SEQUENCE 402 AA; 46041 MW; 2B4A8708C885667C CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 6
Q8VSK3 ID Q8VSK3 PRELIMINARY; PRT; 402 AA.
AC Q8VSK3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Transposase.
GN CP0084.
OS Shigella flexneri 2a.
OG Plasmid pCP301.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=42897;
RN [1]

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 7
Q9K388 ID Q9K388 PRELIMINARY; PRT; 402 AA.
AC Q9K388;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative IS10 transposase.
GN R0076.
OS Salmonella typhi.
OG Plasmid R27.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grotbeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=20247426; PubMed=10783303;
RA Lawley T.D., Burland V.D., Taylor D.E.;
RT "Analysis of the complete nucleotide sequence of the tetracycline-
RT resistance transposon Tn10.";
RL Plasmid 43:235-239(2000).
DR EMBL; AF250878; AAF69914.1; -.
DR EMBL; AF223162; AAF68930.1; -.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Plasmid.
SQ SEQUENCE 402 AA; 46059 MW; 25854FABD793130F CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 8
Q8RQK7 ID Q8RQK7 PRELIMINARY; PRT; 402 AA.
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=301;
RA Jin Q., Zhang J.Y., Liu H., Yang J., Yang F., Zhang X.B., Wang J.H.,
RA Yang G.W., Wu H.T., Dong J., Sun L.L., Xue Y., Zhao A.L., Gao Y.S.,
RA Zhu J.P., Kan B., Chen S.X., Yao Z.J., He B.K., Chen R.S., Ma D.L.,
RA Yuan Z.H., Xu J.G., Wang Y., Shen Y., Lu W.C., Qiang B.Q., Wen Y.M.,
RA Hou Y.D.;
RT "Complete DNA sequence and analysis of the large virulence plasmid
RT pCP301 of Shigella flexneri.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF386526; AAL72416.1; -.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Plasmid.
SQ SEQUENCE 402 AA; 46046 MW; 1E9391EE324A12B6 CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 9
Q9K388 ID Q9K388 PRELIMINARY; PRT; 402 AA.
AC Q9K388;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative IS10 transposase.
GN R0076.
OS Salmonella typhi.
OG Plasmid R27.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grotbeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=20247426; PubMed=10783303;
RA Lawley T.D., Burland V.D., Taylor D.E.;
RT "Analysis of the complete nucleotide sequence of the tetracycline-
RT resistance transposon Tn10.";
RL Plasmid 43:235-239(2000).
DR EMBL; AF250878; AAF69914.1; -.
DR EMBL; AF223162; AAF68930.1; -.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Plasmid.
SQ SEQUENCE 402 AA; 46059 MW; 25854FABD793130F CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 10
Q8RQK7 ID Q8RQK7 PRELIMINARY; PRT; 402 AA.
```


AC Q8RQK7;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Transposase.
GN TRPS.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tim-r01;
RA Nishizawa M., Itoi Y., Ito S., Inoue M.;
RT "Genes expressed in Bacillus cereus.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083420; BAB88926.1; -;
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
SQ SEQUENCE 402 AA; 46070 MW; 53D98F9BD09D5962 CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
DB 247 THCHPS 253

RESULT 9

Q56368
ID Q56368 PRELIMINARY; PRT; 402 AA.
AC Q56368; Q9R2K0;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein (Transposase of TN10) (Tnpr protein).
GN RMPI OR YEDA OR TNPR.
OS Escherichia coli,
OS Methylobacterium aminofaciens,
OS Shigella flexneri, and
OS Salmonella typhimurium.
OG Plasmid R64.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 46896, 623, 602;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; TRANSPOSON=Tn10;
RX MEDLINE=82222187; PubMed=6283536;
RA Halling S.M., Simons R.W., Way J.C., Walsh R.B., Kleckner N.;
RT "DNA sequence organization of IS10-right of Tn10 and comparison with
RT IS10-left.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2608-2612(1982).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; TRANSPOSON=Tn10;
RA Errade P.;
RL Submitted (APR-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.aminofaciens;
RX MEDLINE=99346693; PubMed=10418139;
RA Sakai Y., Mitsui R., Katayama Y., Yanase H., Kato N.;
RT "Organization of the genes involved in the ribulose monophosphate
RT pathway in an obligate methylotrophic bacterium, Methylobacter
RT aminofaciens 77a.";
RL FEMS Microbiol. Lett. 176:125-130(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; TRANSPOSON=Tn10;
RA Chalmers R., Crellin P., Sewitz S., Liplow K.;
RT "The complete nucleotide sequence of transposon Tn10.";

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; PLASMID=R100;
RA Sampei G., Mizobuchi K.;
RT "Organization and diversification of plasmid genomes: complete
RT nucleotide sequence of the R100 genome.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RA Sampei G., Komano T., Sasaki T., Tachibana K., Furuya N., Saito Y.,
RA Suzuki T., Mizobuchi K.;
RT "Organization and diversification of plasmid genomes: complete
RT nucleotide sequence of R64 genome.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=87146423; PubMed=3029698;
RA Komano T., Kubo A., Nishioka T.;
RT "Shufflon: multi-inversion of four contiguous DNA segments of plasmid
RT R64 creates seven different open reading frames.";
RL Nucleic Acids Res. 15:1165-1172(1987).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=89127142; PubMed=3065610;
RA Kubo A., Kusakawa A., Komano T.;
RT "Nucleotide sequence of the rci gene encoding shufflon-specific DNA
RT recombinase in the IncII plasmid R64: homology to the site-specific
RT recombinases of integrase family.";
RL Mol. Gen. Genet. 213:30-35(1988).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=88314948; PubMed=3045094;
RA Komano T., Toyoshima A., Morita K., Nishioka T.;
RT "Cloning and nucleotide sequence of the orit region of the IncII
RT plasmid R64.";
RL J. Bacteriol. 170:4385-4387(1988).
RN [10]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=9117811; PubMed=1848841;
RA Furuya N., Nishioka T., Komano T.;
RT "Nucleotide sequence and functions of the orit operon in IncII plasmid
RT R64.";
RL J. Bacteriol. 173:2231-2237(1991).
RN [11]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=92011438; PubMed=1917882;
RA Furuya N., Komano T.;
RT "Determination of the nick site at orit of IncII plasmid R64: global
RT similarity of orit structures of IncII and IncP plasmids.";
RL J. Bacteriol. 173:6612-6617(1991).
RN [12]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=93352408; PubMed=8349545;
RA Kim S.R., Funayama N., Komano T.;
RT "Nucleotide sequence and characterization of the traABCD region of
RT IncII plasmid R64.";
RL J. Bacteriol. 175:5035-5042(1993).
RN [13]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=95083745; PubMed=7991676;
RA Furuya N., Komano T.;
RT "Surface exclusion gene of IncII plasmid R64: nucleotide sequence and
RT analysis of deletion mutants.";
RL Plasmid 32:80-84(1994).

[14]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhimurium; PLASMID=R64;
RC MEDLINE=96198148; PubMed=8626273;
RX Furuya N., Komano T.;
RA "Nucleotide sequence and characterization of trbABC region of the
RT Inci11 plasmid R64: existence of the pnd gene for plasmid maintenance
RT within the transfer region."
RL J. Bacteriol. 178:1491-1497 (1996).
RN [15]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhimurium; PLASMID=R64;
RC MEDLINE=97428559; PubMed=9281491;
RX Narahara K., Rahman E., Furuya N., Komano T.;
RA "Requirement of a limited segment of the sog gene for plasmid R64
RT conjugation."
RL Plasmid 38:1-11 (1997).
RN [16]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhimurium; PLASMID=R64;
RC MEDLINE=97315231; PubMed=9171405;
RX Kim S.R., Komano T.;
RA "The plasmid R64 thin pilus identified as a type IV pilus."
RT J. Bacteriol. 179:3594-3603 (1997).
RN [17]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhimurium; PLASMID=R64;
RC MEDLINE=98053841; PubMed=9393692;
RX Furuya N., Komano T.;
RA "Mutational analysis of the R64 orit region: requirement for precise
RT location of the Nika-binding sequence."
RL J. Bacteriol. 179:7291-7297 (1997).
RN [18]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhimurium; PLASMID=R64;
RC MEDLINE=98268996; PubMed=9603870;
RX Yoshida T., Furuya N., Ishikura M., Isobe T., Haino-Fukushima K.,
RA Ogawa T., Komano T.;
RT "Purification and characterization of thin pili of Inci11 plasmids
RT Colib-P9 and R64: formation of PiliV-specific cell aggregates by type
RT IV pili."
RL J. Bacteriol. 180:2842-2848 (1998).
RN [19]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhimurium; PLASMID=R64;
RC MEDLINE=20223621; PubMed=10760136;
RX Komano T., Yoshida T., Narahara K., Furuya N.;
RA "The transfer region of Inci11 plasmid R64: similarities between R64
RT tra genes and Legionella icm/dot genes."
RL Mol. Microbiol. 35:1348-1359 (2000).
DR EMBL; J01829; AAA88660.1; -;
DR EMBL; AB026428; BAA83097.1; -;
DR EMBL; AF162223; AAD50250.1; -;
DR EMBL; AP000342; BAA78838.1; -;
DR EMBL; AP005147; BAB91573.1; -;
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
DB 247 THCHPS 253

RESULT 10
Q8GC95
ID Q8GC95 PRELIMINARY; PRT; 402 AA.
AC Q8GC95;

01-MAR-2003 (TREMBLrel. 23, Created)
01-MAR-2003 (TREMBLrel. 23, Last sequence update)
01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Transposase.
GN TNP.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3009;
RA Celschlaeger T.A.;
RT "Cloning, sequencing and expression of an invasion determinant from
RT Citrobacter freundii strain 3009."
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ508060; CAD56977.1; -;
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
DB 247 THCHPS 253

RESULT 11
Q9HSN6
ID Q9HSN6 PRELIMINARY; PRT; 402 AA.
AC Q9HSN6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ23250.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026903; BAB15586.1; -;
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Hypothetical protein.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 4; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
DB 247 THCHPS 253

RESULT 12
Q9BTF6
ID Q9BTF6 PRELIMINARY; PRT; 402 AA.
AC Q9BTF6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC004131; AA04131.1; -;
 DR InterPro; IPR002559; Transposase_11.
 DR Pfam; PF01609; Transposase_11; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 402 AA; 46041 MW; 2B4A8708C885667C CRC64;

Query Match 88.6%; Score 39; DB 4; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
 |||||
 Db 247 THCHPS 253

RESULT 13
 Q95T96
 ID Q95T96 PRELIMINARY; PRT; 402 AA.
 AC Q95T96;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE GM13045p.
 GN CG4800.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY058482; AAL13711.1; -;
 DR FlyBase; FBgn0037874; CG4800.
 DR InterPro; IPR002559; Transposase_11.
 DR Pfam; PF01609; Transposase_11; 1.
 SQ SEQUENCE 402 AA; 46041 MW; 2B4A8708C885667C CRC64;

Query Match 88.6%; Score 39; DB 5; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
 |||||
 Db 247 THCHPS 253

RESULT 14
 Q95T96
 ID Q95T96 PRELIMINARY; PRT; 402 AA.
 AC Q95T96;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE GH10639p.
 GN CG8216.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY060273; AAL25312.1; -;
 DR FlyBase; FBgn0033358; CG8216.
 DR InterPro; IPR002559; Transposase_11.
 DR Pfam; PF01609; Transposase_11; 1.
 SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 5; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
 |||||
 Db 247 THCHPS 253

RESULT 15
 Q8QTC0
 ID Q8QTC0 PRELIMINARY; PRT; 402 AA.
 AC Q8QTC0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE WSSV410.
 OS White spot syndrome virus (WSSV).
 OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
 OX NCBI_TaxID=92652;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 RX MEDLINE=20517548; PubMed=11062040;
 RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
 RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
 RT "Identification and characterization of a shrimp white spot syndrome virus (WSSV) gene that encodes a novel chimeric polypeptide of cellular-type thymidine kinase and thymidylate kinase.";
 RL Virology 277:100-110(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 RX MEDLINE=21844071; PubMed=11853398;
 RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
 RA Lo C.F., Kou G.H.;
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white spot syndrome virus and characterization of the motif important for targeting VP35 to the nuclei of transfected insect cells.";
 RL Virology 293:44-53(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Taiwan;
 RA Lo C.-F., Kou G.-H.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF440570; AAL89278.1; -;
 DR InterPro; IPR002559; Transposase_11.
 DR Pfam; PF01609; Transposase_11; 1.
 SQ SEQUENCE 402 AA; 46073 MW; 2B4A8708D7930BAC CRC64;

Query Match 88.6%; Score 39; DB 12; Length 402;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
 |||||
 Db 247 THCHPS 253

Search completed: February 11, 2004, 17:09:46
Job time : 26.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 MMMNRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	70	Q9LZD2	Q9LZD2 arabidopsis
2	30	88.2	192	Q9SQN9	Q9sqn9 arabidopsis
3	30	88.2	249	Q8LDG1	Q8ldg1 arabidopsis
4	30	88.2	249	Q9SVL0	Q9svl0 arabidopsis
5	30	88.2	353	Q9STS7	Q9sts7 arabidopsis
6	30	88.2	364	Q9AWZ8	Q9awz8 oryza sativ
7	28	82.4	129	Q95RT0	Q95rt0 drosophila
8	28	82.4	307	Q8WA93	Q8wa93 narceus ann
9	28	82.4	307	Q8ZM25	Q8zm25 salmonella
10	27	79.4	39	Q00868	Q00868 plasmodium
11	27	79.4	70	Q8LE23	Q8le23 arabidopsis
12	27	79.4	103	Q8H4V8	Q8h4v8 oryza sativ
13	27	79.4	147	Q25870	Q25870 plasmodium
14	27	79.4	176	Q9WBQ4	Q9wbq4 apis labori
15	27	79.4	238	Q9FKE0	Q9fke0 arabidopsis
16	27	79.4	283	Q9IQD9	Q9iqd9 drosophila

17	27	79.4	309	10	Q8RUX0	Q8rux0 oryza sativ
18	27	79.4	326	8	Q8M355	Q8m355 saccharomyc
19	27	79.4	365	2	Q9KH45	Q9kh45 pantoea agg
20	27	79.4	427	10	Q9S7L2	Q9s7l2 arabidopsis
21	27	79.4	501	5	Q9NEX5	Q9nbx5 drosophila
22	27	79.4	747	10	Q8LHX4	Q8lhx4 oryza sativ
23	27	79.4	1070	5	Q8IEC2	Q8iec2 plasmodium
24	27	79.4	2144	11	Q8R2J0	Q8r2j0 mus musculu
25	27	79.4	3013	5	Q8IJZ5	Q8ijz5 plasmodium
26	26	76.5	60	4	Q9H3F1	Q9h3f1 homo sapien
27	26	76.5	61	4	Q8N6W1	Q8n6w1 homo sapien
28	26	76.5	213	16	Q8PBI8	Q8pbis xanthomonas
29	26	76.5	215	16	Q8PN41	Q8pn41 xanthomonas
30	26	76.5	216	16	Q9PDQ3	Q9pdq3 xylella fas
31	26	76.5	225	10	Q9LJV9	Q9ljv9 arabidopsis
32	26	76.5	262	8	Q9MHI0	Q9mhi0 melitaea pe
33	26	76.5	302	10	Q9M9E6	Q9m9e6 arabidopsis
34	26	76.5	309	16	Q9K953	Q9k953 bacillus ha
35	26	76.5	312	11	Q8VGM0	Q8vgm0 mus musculu
36	26	76.5	354	10	Q9SRQ0	Q9srq0 arabidopsis
37	26	76.5	374	5	Q9NGS0	Q9ngs0 bombyx mori
38	26	76.5	375	10	Q9SMR3	Q9smr3 arabidopsis
39	26	76.5	380	5	Q9NFT5	Q9nft5 anopheles g
40	26	76.5	380	5	Q8WSX8	Q8wsx8 anopheles g
41	26	76.5	382	2	Q9FCY8	Q9fcy8 erwinia ste
42	26	76.5	436	10	Q9C954	Q9c954 arabidopsis
43	26	76.5	437	5	O45007	O45007 caenorhabdi
44	26	76.5	484	16	Q92YF2	Q92yf2 rhizobium m
45	26	76.5	509	8	Q9GCJ9	Q9gcj9 parides pho

ALIGNMENTS

RESULT 1

Q9LZD2
ID Q9LZD2 PRELIMINARY; PRT; 70 AA.
AC Q9LZD2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 8.0 kDa protein.
GN F12E4_330.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RP Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RP Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene F12E4_330 (GI:7378640).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AL162751; CAB83316.1; -
DR EMBL; AY040018; AAK64175.1; -
KW Hypothetical protein.
SQ SEQUENCE 70 AA; 8024 MW; A224B96F68370BC7 CRC64;

Query Match 94.1%; Score 32; DB 10; Length 70;
Best Local Similarity 85.7%; Pred. No. 1.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
| | | | |
Db 3 MMTMMRI 9

RESULT 2

Q9SQN9 PRELIMINARY; PRT; 192 AA.
AC Q9SQN9; SEQUENCE FROM N.A.
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 21.2 kDa protein.
GN F19K16.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome 1 BAC F19K16 genomic sequence";
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL
RN [2]
RP SEQUENCE FROM N.A.
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F20B17 from chromosome
I.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011717; AAG52257.1; -.
DR EMBL; AC010793; AAF68112.1; -.
DR HSSP; P29602; 1JER.
DR InterPro; IPR003245; Plcyanin like.
DR Pfam; PF02298; Cu_bind_like; 1.
DR ProDom; PD003122; plcyanin_like; 1.
KW Hypothetical protein.
SQ SEQUENCE 192 AA; 21197 MW; 2322FD669A95A04C CRC64;

Query Match 88.2%; Score 30; DB 10; Length 192;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
| | | | |
Db 2 MMTMMR 7

RESULT 3

Q8LDG1 PRELIMINARY; PRT; 249 AA.
AC Q8LDG1;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troughan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY086020; AAM63229.1; -.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR006455; Homeo_ZF_HD.
DR InterPro; IPR006456; ZF_HD_N.
DR TIGRFAMS; TIGR01565; homeo_ZF_HD; 1.
DR TIGRFAMS; TIGR01566; ZF_HD_prot_N; 1.
DR PROSITE; PS50071; HOMEODOM_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 249 AA; 28605 MW; 4CD673D6EEE665A7 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 249;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
| | | | |
Db 174 MMTMMR 179

RESULT 4

Q9SVL0 PRELIMINARY; PRT; 249 AA.
AC Q9SVL0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical 28.7 kDa protein.
GN F18B3.170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,
RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C., Quetier F.,
RA Salanoubat M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049862; CAB42918.1; -.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR006455; Homeo_ZF_HD.
DR InterPro; IPR006456; ZF_HD_N.
DR InterPro; IPR006857; ZnF_HD_dimer.

DR Pfam; PF04770; ZF-HD dimer; 1.
DR TIGRFAMS; TIGR01565; homeo ZF HD; 1.
DR TIGRFAMS; TIGR01566; ZF HD_prot_N; 1.
DR PROSITE; PS0071; HOMEBOX_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 249 AA; 28658 MW; F0C163A282D51BB4 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 249;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
DB 174 MMTMMR 179

RESULT 5
Q9STS7 PRELIMINARY; PRT; 353 AA.
AC Q9STS7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical 39.1 kDa protein.
GN T23J7.190.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RA Choisne N., Robert C., Brottier P., Wincker P., Cattolico L.,
RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Mayer K.F.X.,
RA Lemcke K., Schueller C., Quetier F., Salanoubat M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,
RA Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
RA Yamada K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene T23J7.190/AT3g47860 (GI:4741203).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049746; CAB41869.1; -
DR EMBL; AY044334; AAK73275.1; -
DR EMBL; AY035165; AAK59669.1; -
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR00566; Lipocln_cytFABP.
DR Pfam; PF00061; lipocalin; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 353 AA; 39116 MW; 14A83B0CF1BA8D0D CRC64;

Query Match 88.2%; Score 30; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
DB 122 MMTMMR 127

RESULT 6
Q9AWZ8 PRELIMINARY; PRT; 364 AA.
AC Q9AWZ8;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 3-methyl-2-oxobutanate hydroxy-methyl-transferase-like protein.
GN P0492F05.14 OR P0443E07.6.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530, 39947;
[1]
SEQUENCE FROM N.A.
RC SPECIES=O.sativa; STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0492F05.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RC SPECIES=O.sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0443E07.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002902; BAB32712.1; -
DR EMBL; AP002900; BAB92103.1; -
DR Gramene; Q9AWZ8;
DR InterPro; IPR003700; Pantoate transf.
DR Pfam; PF02548; Pantoate_transf; 1.
DR TIGRFAMS; TIGR00222; panB; 1.
KW Transferase.
SQ SEQUENCE 364 AA; 37920 MW; 3393C0A29F2B1512 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 364;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
DB 1 MMTMMR 6

RESULT 7
Q95RT0 PRELIMINARY; PRT; 129 AA.
AC Q95RT0;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE LD12764p.
GN BCDNA:LD12764.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunco J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,

RA Yu C., Lewis S.B., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061159; AAL28707.1; -;
DR FlyBase; FBgn0047212; BcDNA:LD12764.
SQ SEQUENCE 129 AA; 14240 MW; C349AABE9EC3DBEF CRC64;

Query Match 82.4%; Score 28; DB 5; Length 129;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
|||||
Db 13 MMTMMTL 19

RESULT 8
Q8WA93 PRELIMINARY; PRT; 307 AA.
AC Q8WA93;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
oxidoreductase chain 1).
GN NAD1.
OS Narceus annularus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Diplopoda; Helminthomorpha;
OC Spirobolidae; Narceus.
OX NCBI_TaxID=174156;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21660208; PubMed=11801744;
RA Lavrov D.V., Boore J.L., Brown W.M.;
RT "Complete mtDNA Sequences of Two Millipedes Suggest a New Model for
RT Mitochondrial Gene Rearrangements: Duplication and Nonrandom Loss.";
RL Mol. Biol. Evol. 19:163-169(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Lavrov D.V.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
DR EMBL; AY055727; AAL18214.1; -;
DR InterPro; IPR001694; Resp_NADH_dh1.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
SQ SEQUENCE 307 AA; 34669 MW; 03F837A142D8A24D CRC64;

Query Match 82.4%; Score 28; DB 8; Length 307;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
|||||
Db 233 MMTMMSL 239

RESULT 9
Q8ZM25 PRELIMINARY; PRT; 307 AA.
AC Q8ZM25;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative xylanase/chitin deacetylase.
GN STM3132.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;

Query Match 82.4%; Score 28; DB 8; Length 307;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008844; AAL22006.1; -;
DR InterPro; IPR002509; Polysac_deacet.
DR Pfam; PF01522; Polysacc_deacet; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Hypothetical protein;
KW Complete proteome.
SQ SEQUENCE 307 AA; 35194 MW; B14436094A935245 CRC64;

Query Match 82.4%; Score 28; DB 16; Length 307;
Best Local Similarity 71.4%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
|||||
Db 249 MMTMLRL 255

RESULT 10
O00868 PRELIMINARY; PRT; 39 AA.
ID O00868
AC O00868;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE HMG-like protein Pf16 (Fragment).
GN PF16.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB3;
RX MEDLINE=92118046; PubMed=1731798;
RA Guntaka R.V., Kandala J.C., Reddy V.D.;
RT "Cloning and characterization of a highly conserved HMG-like protein
RT (PF16) gene from Plasmodium falciparum.";
RL Biochem. Biophys. Res. Commun. 182:412-419(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HB3;
RA Nambiar A., Kandala J.C., Dolan S.A., Jensen J.J., Guntaka R.V.;
RT "Molecular cloning and characterization of a cDNA for the highly
RT conserved HMG-like protein (Pf16) gene of Plasmodium falciparum.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U97128; AAB58116.1; -;
FT NON_TER 39
SQ SEQUENCE 39 AA; 5037 MW; 012D80DD1F566823 CRC64;

Query Match 79.4%; Score 27; DB 5; Length 39;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
|||||
Db 21 MMTMMK 26

RESULT 11
Q8LE23 PRELIMINARY; PRT; 70 AA.
ID Q8LE23
AC Q8LE23;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN Arabidopsis thaliana (Mouse-ear cross).
OS Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troupkhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troupkhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085662; AAM67306.1; -
KW Hypothetical protein.
SQ SEQUENCE 70 AA; 7270 MW; 10C0764E0986E031 CRC64;

Query Match 79.4%; Score 27; DB 10; Length 70;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
Db 1 MMTMMK 6

RESULT 12

Q8H4V8 ID Q8H4V8 PRELIMINARY; PRT; 103 AA.
AC Q8H4V8;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE OJ1484_G09.19 protein.
GN OJ1484_G09.19.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT clone:OJ1484_G09.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003913; BAC24906.1; -
SQ SEQUENCE 103 AA; 11753 MW; 22AB780BBF89FCB0 CRC64;

Query Match 79.4%; Score 27; DB 10; Length 103;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 7
Db 45 MMTMMK 51

RESULT 13

Q25870 ID Q25870 PRELIMINARY; PRT; 147 AA.
AC Q25870;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE High mobility group-like protein.
GN PS16.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92118046; PubMed=1731798;
RA Guntaka R.V., Kandala J.C., Vudem D.;
RT "Cloning and characterization of a highly conserved HMG-like protein
RT (PF16) gene from Plasmodium falciparum.";
RL Biochem. Biophys. Res. Commun. 182:412-419(1992).
DR EMBL; M86518; AAA29615.1; -
SQ SEQUENCE 147 AA; 17175 MW; A46C5D47A62DF75A CRC64;

Query Match 79.4%; Score 27; DB 5; Length 147;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
Db 11 MMTMMK 16

RESULT 14

Q8WBQ4 ID Q8WBQ4 PRELIMINARY; PRT; 176 AA.
AC Q8WBQ4;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
GN ND4.
OS Apis laboriosa (Himalayan honeybee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=183418;
RN [1]
RP SEQUENCE FROM N.A.
RA Leelamanit W., Amano K.;
RT "The NADH Dehydrogenases of Apis mellifera, A. cerana, A. dorsata, A.
RT laboriosa and A. florea: Sequence Comparison and Genetic Diversity.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF447064; AAL60208.1; -
DR InterPro; IPR003918; NADH_oxred4.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
DR PRINTS; PR01437; NUOXDRDTASE4.
KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT NON_TER 1 176
FT NON_TER 176 176
SQ SEQUENCE 176 AA; 19981 MW; 9E405EED9CF06687 CRC64;

Query Match 79.4%; Score 27; DB 8; Length 176;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
Db 34 MMTMMK 39

RESULT 15

Q9FKE0 ID Q9FKE0 PRELIMINARY; PRT; 238 AA.
AC Q9FKE0;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Gb|AAF30306.1.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones."
RL DNA Res. 5:203-216(1998).
DR EMBL; AB012241; BAB09032.1; -.
DR InterPro; IPR006121; HeavyMe.transpt.
DR InterPro; IPR006191; Metal_bind.
DR Pfam; PF00403; HMA; 1.
DR PROSITE; PS50846; HMA 2; 1.
SQ SEQUENCE 238 AA; 27717 MW; F2F486C355011632 CRC64;

Query Match 79.4%; Score 27; DB 10; Length 238;
Best Local Similarity 71.4%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
Db 103 M M M M M M M M 109

Search completed: February 11, 2004, 17:09:48
Job time : 26.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 MNNMRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	85.3	132	1 F802 SCHMA	P16464 schistosoma
2	27	79.4	146	1 2SS BEREX	P04403 bertholleti
3	27	79.4	1610	1 CCAD MESAU	Q99244 mesocricetu
4	27	79.4	2161	1 CCAD HUMAN	Q01668 homo sapien
5	27	79.4	2203	1 CCAD RAT	P27732 rattus norv
6	26	76.5	395	1 RIF2 YEAST	Q06208 saccharomyc
7	26	76.5	403	1 HRP1 ERWAM	Q01099 erwinia amy
8	26	76.5	438	1 POP1 CAEEL	Q10666 caenorhabdi
9	26	76.5	875	1 AMD2 XENLA	P12890 xenopus lae
10	25	73.5	313	1 NULM RHISA	Q99824 rhipicephal
11	25	73.5	321	1 AOX1 SOYBN	Q07185 glycine max
12	25	73.5	325	1 CY1 HUMAN	P08574 homo sapien
13	25	73.5	355	1 HKL6 LYCES	O22299 lycopersico
14	25	73.5	377	1 Y4WD RHISN	P55682 rhizobium s
15	25	73.5	382	1 STM ARATH	Q38874 arabidopsis
16	25	73.5	383	1 STM BRAOL	Q9m6d9 brassica ol
17	25	73.5	400	1 AMD1 XENLA	P08478 xenopus lae
18	25	73.5	457	1 NORM ECO57	P58164 escherichia
19	25	73.5	457	1 NORM ECOLI	P37340 escherichia
20	25	73.5	554	1 NUSM APILI	P34855 apis mellif
21	25	73.5	783	1 YKR2 CAEEL	P34308 caenorhabdi
22	25	73.5	931	1 DING BACSU	P54394 bacillus su
23	24	70.6	96	1 CLV3 ARATH	Q9xf04 arabidopsis
24	24	70.6	99	1 PD11 HUMAN	Q9y6g1 homo sapien
25	24	70.6	99	1 PD11 MOUSE	P56983 mus musculu
26	24	70.6	179	1 Y281 METJA	Q57729 methanococc
27	24	70.6	181	1 GC12 PSEAE	Q9i351 pseudomonas
28	24	70.6	186	1 GC11 PSEAE	Q9hyg8 pseudomonas
29	24	70.6	265	1 CTRC NEIMA	P57012 neisseria m
30	24	70.6	322	1 PYRB_XYLFA	Q9pbb8 xylella fas
31	24	70.6	423	1 MB12 YEAST	P03873 saccharomyc
32	24	70.6	434	1 AVT CATCO	Q90352 catostomus
33	24	70.6	437	1 PUR2_XYLFA	Q9pc09 xylella fas

34	24	70.6	462	1 ZRAS KLEOX	Q9ape0 klebsiella
35	24	70.6	503	1 CP51 HUMAN	Q16850 homo sapien
36	24	70.6	562	1 ILVD PYRAE	Q8zyu6 pyrobaculum
37	24	70.6	582	1 YC73 MYCTU	Q11046 mycobacteri
38	24	70.6	687	1 SRB4 YEAST	P32569 saccharomyc
39	24	70.6	902	1 MOPT ARATH	P93024 arabidopsis
40	24	70.6	1579	1 SSK2 YEAST	P53599 saccharomyc
41	23	67.6	49	1 Y495 TREPA	O83508 treponema p
42	23	67.6	72	1 YNEF_BACSU	P45708 bacillus su
43	23	67.6	75	1 YA26 STAEF	Q8csp3 staphylococ
44	23	67.6	80	1 YD43 STAAE	Q99ud3 staphylococ
45	23	67.6	89	1 Y703 ARCFU	O29555 archaeoglob

ALIGNMENTS

RESULT 1

F802 SCHMA
ID_F802 SCHMA STANDARD; PRT; 132 AA.
AC P16464;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Female specific 800 protein (FS800).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Puerto Rican;
RX MEDLINE=89181810; PubMed=2927441;
RA Reis M.G., Kuhn J., Blanton R., Davis A.H.;
RT "Localization and pattern of expression of a female specific mRNA in
Schistosoma mansoni";
RL Mol. Biochem. Parasitol. 32:113-119(1989).
CC -!- FUNCTION: FS800 is likely to have some function in the production
or maintenance of the schistosome egg. It may have a function
unrelated to eggshell formation.
CC -!- DEVELOPMENTAL STAGE: Highest level only in mature worms, i.e.,
during egg production.
CC -!- MISCELLANEOUS: The two F800 proteins are read from two
overlapping reading frames.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; J03999; AAA29884.1; - ILE-RICH.
FT DOMAIN 2 21
FT DOMAIN 47 53 POLY-MET.
SQ SEQUENCE 132 AA; 15561 MW; 02C77F42A25E120E CRC64;

Query Match 85.3%; Score 29; DB 1; Length 132;
Best Local Similarity 71.4%; Pred. No. 2;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNMRL 7
Db 49 MNNMRL 55

RESULT 2

2SS BEREX STANDARD; PRT; 146 AA.
ID_2SS BEREX
AC P04403; P04402;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 2S sulfur-rich seed storage protein precursor (Allergen Ber e 1).
GN BE2S1 AND BE2S2.
OS Bertholletia excelsa (Brazil nut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Ericales; Lecythidaceae; Bertholletia.
OX NCBI_TaxID=3645;
RN [1]
RN SEQUENCE FROM N.A.
RA Altenbach S.B., Pearson K.W., Leung F.W., Sun S.S.M.;
RT "Cloning and sequence analysis of a cDNA encoding a Brazil nut protein
RT exceptionally rich in methionine."
RL Plant Mol. Biol. 8:239-250(1987).
RN [2]
RN SEQUENCE FROM N.A.
RA Bassuener R.;
RL Submitted (DEC-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=91370890; PubMed=1840683;
RA Gander E.S., Holmstroem K.O., de Paiva G.R., de Castro L.A.B.,
RA Carneiro M., Grossi de Sa M.F.;
RT "Isolation, characterization and expression of a gene coding for a 2S
RT albumin from Bertholletia excelsa (Brazil nut).";
RL Plant Mol. Biol. 16:437-448(1991).
RN [4]
RN SEQUENCE OF 37-64 AND 70-142.
RX MEDLINE=87004679; PubMed=3758080;
RA Ampe C., van Damme J., de Castro L.A.B., Sampaio M.J.A.M.,
RA van Montagu M., Vandekerckhove J.;
RT "The amino-acid sequence of the 2S sulphur-rich proteins from seeds
RT of Brazil nut (Bertholletia excelsa H.B.K.).";
RL Eur. J. Biochem. 159:597-604(1986).
CC -!- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC -!- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC LINKED BY DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17146; AAA33010.1; -
DR EMBL; X57027; CAA40343.1; -
DR EMBL; X57028; CAA40344.1; -
DR EMBL; X54490; CAA38362.1; -
DR EMBL; X54491; CAA38363.1; ALT_SEQ.
DR EMBL; A13818; CAA01131.1; -
DR PIR; A25802; A25802.
DR PIR; S14946; S14946.
DR PDB; 1GYS; 30-APR-02.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR00617; Napin.
DR Pfam; PF00234; trypan_alpha_aml; 1.
DR PRINTS; PR00496; NAPIN.
DR SMART; SM00499; AAI; 1.
KW Seed storage protein; Signal; Allergen; Pyrrolidone carboxylic acid;
KW 3D-structure.
FT SIGNAL 1 22
FT PROPEP 23 36
FT CHAIN 37 64 SMALL CHAIN.
FT PROPEP 65 69
FT CHAIN 70 142 LARGE CHAIN 1B.
FT PROPEP 143 146
FT MOD RES 37 37 PYRROLIDONE CARBOXYLIC ACID.
FT VARIANT 91 91 S -> E (IN VARIANT 1A).
FT CONFLICT 38 39 EE -> QQ (IN REF. 4).
FT CONFLICT 122 122 L -> M (IN REF. 4).

FT CONFLICT 126 126 I -> L (IN REF. 4).
SQ SEQUENCE 146 AA; 16911 MW; A7DF778FD766410D CRC64;
Query Match 79.4%; Score 27; DB 1; Length 146;
Best Local Similarity 83.3%; Pred. No. 6.7;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 MMMMRL 7
Db 99 MMTMRM 104
RESULT 3
CCAD_MESAU STANDARD; PRT; 1610 AA.
ID CCAD_MESAU Q99245;
AC Q99244; Q99245;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1D subunit (Calcium
DE channel, L type, alpha-1 polypeptide isoform 2).
GN CACNA1D OR CACNL1A2 OR CCHL1A2 OR CACH3 OR CACNA4.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM HCA3A).
RP TISSUE=Insulinoma;
RX MEDLINE=93149124; PubMed=1337146;
RA Yaney G.C., Wheeler M.B., Wei X., Perez-Reyes E., Birnbaumer L.,
RA Boyd A.E. III, Moss L.G.;
RT "Cloning of a novel alpha 1-subunit of the voltage-dependent calcium
RT channel from the beta-cell."
RL Mol. Endocrinol. 6:2143-2152(1992).
RN [2]
RN SEQUENCE OF 1146-1441 FROM N.A. (ISOFORMS CACH3B AND CACH3D).
RP TISSUE=Heart;
RX MEDLINE=91056091; PubMed=2173707;
RA Perez-Reyes E., Wei X., Castellano A., Birnbaumer L.;
RT "Molecular diversity of L-type calcium channels. Evidence for
RT alternative splicing of the transcripts of three non-allelic genes."
RL J. Biol. Chem. 265:20430-20436(1990).
CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
CC PHENYLALKYLAMINES, BENZOTHIAZEPINES, AND BY OMEGA-AGATOXIN-III
CC (OMEGA-AGA-IIIa). THEY ARE HOWEVER INSENSITIVE TO OMEGA-AGATOXIN-
CC GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=HCA3A;
CC IsoId=Q99244-1; Sequence=Displayed;
CC Name=CACH3B;
CC IsoId=Q99244-3; Sequence=Not described;
CC Name=CACH3D;
CC IsoId=Q99244-2; Sequence=VSP_000915;
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART AND SKELETAL MUSCLE.

CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M57969; AAB59702.1; -.
CC EMBL; M57970; AAA62807.1; -.
CC PIR; A46227; A46227.
CC InterPro; IPR001682; Ca/Na_pore.
CC InterPro; IPR002077; Ca_channel.
CC InterPro; IPR002111; Cat_channel_TrpL.
CC InterPro; IPR005821; Ion_trans.
CC InterPro; IPR005446; LVDCCALPHAL.
CC InterPro; IPR005452; LVDCCALPHALD.
CC InterPro; IPR005820; M+channel_nlg.
CC Pfam; PF00520; ion_trans; 4.
CC PRINTS; PR00167; CACHANNEL.
CC PRINTS; PR01630; LVDCCALPHAL.
CC PRINTS; PR01636; LVDCCALPHALD.
CC PRINTS; PR01636; LVDCCALPHALD.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel; Glycoprotein; Repeat; Multigene family;
KW Calcium-binding; Phosphorylation; Alternative splicing.
FT REPEAT 112 408
FT REPEAT 508 754
FT REPEAT 871 1153
FT REPEAT 1190 1465
FT DOMAIN 1 125
FT TRANSMEM 126 144
FT DOMAIN 145 162
FT TRANSMEM 163 182
FT DOMAIN 183 194
FT TRANSMEM 195 213
FT DOMAIN 214 234
FT TRANSMEM 235 253
FT DOMAIN 254 272
FT TRANSMEM 273 292
FT DOMAIN 293 380
FT TRANSMEM 381 405
FT DOMAIN 406 522
FT TRANSMEM 523 542
FT DOMAIN 543 557
FT TRANSMEM 558 576
FT DOMAIN 577 584
FT TRANSMEM 585 603
FT DOMAIN 604 613
FT TRANSMEM 614 632
FT DOMAIN 633 651
FT TRANSMEM 652 672
FT DOMAIN 673 726
FT TRANSMEM 727 751
FT DOMAIN 752 884
FT TRANSMEM 885 903
FT DOMAIN 904 919
FT TRANSMEM 920 939
FT DOMAIN 940 951
FT TRANSMEM 952 970
FT DOMAIN 971 976
FT TRANSMEM 977 996
FT DOMAIN 997 1015
FT TRANSMEM 1016 1035
FT DOMAIN 1036 1125
FT TRANSMEM 1126 1146

FT DOMAIN 1147 1203
FT TRANSMEM 1204 1222
FT DOMAIN 1223 1237
FT TRANSMEM 1238 1257
FT DOMAIN 1258 1264
FT TRANSMEM 1265 1286
FT DOMAIN 1287 1311
FT TRANSMEM 1312 1331
FT DOMAIN 1332 1350
FT TRANSMEM 1351 1370
FT DOMAIN 1371 1437
FT TRANSMEM 1438 1462
FT DOMAIN 1463 1610
FT DOMAIN 1 6
FT TRANSMEM 652 658
FT DOMAIN 826 836
FT DOMAIN 428 445
FT SITE 363 363
FT SITE 704 704
FT SITE 1099 1099
FT SITE 1404 1404
FT BINDING 1073 1163
FT BINDING 1418 1484
FT BINDING 1430 1473
FT MOD_RES 1473 1473
FT CA_BIND 1491 1502
FT CARBOHYD 154 154
FT CARBOHYD 224 224
FT CARBOHYD 328 328
FT VARSPPLIC 1261 1303
FT
FT
SQ SEQUENCE 1610 AA; 182327 MW; B3B2E3794D936F79 CRC64;
/FTid=VSP 000915.
Query Match 79.4%; Score 27; DB 1; Length 1610;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMTMMR 6
Db 2 MMTMMK 7
RESULT 4
CCAD_HUMAN STANDARD; PRT; 2161 AA.
ID CCAD_HUMAN STANDARD; PRT; 2161 AA.
AC Q01668; Q13916; Q13931;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1D subunit (Calcium
DE channel, L type, alpha-1 polypeptide, isoform 2).
GN CACNA1D OR CACNL1A2 OR CCHL1A2 OR CACH3 OR CACN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM NEURONAL-TYPE).
RC TISSUE=Neuroblastoma;
RA MEDLINE=92110010; PubMed=1309651;
RA Williams M.E., Feldman D.H., McCue A.F., Brenner R., Velicelebi G.,
RA Ellis S.B., Harpold M.M.;
RT "Structure and functional expression of alpha 1, alpha 2, and beta
RT subunits of a novel human neuronal calcium channel subtype."
RL Neuron 8:71-84(1992).
RN [2]

RP SEQUENCE FROM N.A. (ISOFORM BETA-CELL-TYPE).

RX TISSUE=Pancreatic islets;

RA MEDLINE=92115705; PubMed=1309948;

RA Seino S., Chen L., Seino M., Blondel O., Takeda J., Johnson J.H.,

RA Bell G.I.;

RT "Cloning of the alpha 1 subunit of a voltage-dependent calcium

RT channel expressed in pancreatic beta cells.;"

RL Proc. Natl. Acad. Sci. U.S.A. 89:584-588(1992).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM BETA-CELL-TYPE), AND VARIANT NIDDM.

RX MEDLINE=96044438; PubMed=7557998;

RA Yamada Y., Masuda K., Li Q., Ihara Y., Kubota A., Miura T.,

RA Nakamura K., Fujii Y., Seino S., Seino Y.;

RT "The structures of the human calcium channel alpha 1 subunit

RT (CACNL1A2) and beta subunit (CACNLB3) genes.;"

RL Genomics 27:312-319(1995).

CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE

CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED

CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE

CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,

CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D

CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)

CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)

CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),

CC PHENYLALKYLAMINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-III

CC (OMEGA-AGA-IIIA). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-

CC GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).

CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT

CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS

CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-

CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS

CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM

CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA

CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Comment=Additional isoforms seem to exist;

CC Name=Neuronal-type;

CC IsoId=Q01668-1; Sequence=Displayed;

CC Name=Beta-cell-type;

CC IsoId=Q01668-2; Sequence=VSP 000913, VSP 000914;

CC -!- TISSUE SPECIFICITY: EXPRESSED IN PANCREATIC ISLETS AND IN BRAIN,

CC WHERE IT HAS BEEN SEEN IN HIPPOCAMPUS, BASAL GANGLIA, HABENULA AND

CC THALAMUS. NO EXPRESSION IN SKELETAL MUSCLE.

CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE

CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE

CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS

CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A

CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.

CC -!- POLYMORPHISM: A CHANGE FROM SEVEN TO EIGHT ATG TRINUCLEOTIDE

CC REPEATS, RESULTING IN AN ADDITIONAL N-TERMINAL METHIONINE, HAS

CC BEEN FOUND IN A PATIENT WITH NON-INSULIN-DEPENDENT DIABETES

CC MELLITUS (NIDDM).

CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; M76558; AAA58402.1; -

CC EMBL; M83566; AAA35629.1; -

CC EMBL; D43747; BAA07804.1; -

CC EMBL; D43705; BAA07804.1; JOINED.

CC EMBL; D43706; BAA07804.1; JOINED.

CC EMBL; D43707; BAA07804.1; JOINED.

CC EMBL; D43708; BAA07804.1; JOINED.

CC EMBL; D43709; BAA07804.1; JOINED.

DR EMBL; D43710; BAA07804.1; JOINED.

DR EMBL; D43711; BAA07804.1; JOINED.

DR EMBL; D43712; BAA07804.1; JOINED.

DR EMBL; D43713; BAA07804.1; JOINED.

DR EMBL; D43714; BAA07804.1; JOINED.

DR EMBL; D43715; BAA07804.1; JOINED.

DR EMBL; D43716; BAA07804.1; JOINED.

DR EMBL; D43717; BAA07804.1; JOINED.

DR EMBL; D43718; BAA07804.1; JOINED.

DR EMBL; D43719; BAA07804.1; JOINED.

DR EMBL; D43720; BAA07804.1; JOINED.

DR EMBL; D43721; BAA07804.1; JOINED.

DR EMBL; D43722; BAA07804.1; JOINED.

DR EMBL; D43723; BAA07804.1; JOINED.

DR EMBL; D43724; BAA07804.1; JOINED.

DR EMBL; D43725; BAA07804.1; JOINED.

DR EMBL; D43726; BAA07804.1; JOINED.

DR EMBL; D43727; BAA07804.1; JOINED.

DR EMBL; D43728; BAA07804.1; JOINED.

DR EMBL; D43729; BAA07804.1; JOINED.

DR EMBL; D43730; BAA07804.1; JOINED.

DR EMBL; D43731; BAA07804.1; JOINED.

DR EMBL; D43732; BAA07804.1; JOINED.

DR EMBL; D43733; BAA07804.1; JOINED.

DR EMBL; D43734; BAA07804.1; JOINED.

DR EMBL; D43735; BAA07804.1; JOINED.

DR EMBL; D43736; BAA07804.1; JOINED.

DR EMBL; D43737; BAA07804.1; JOINED.

DR EMBL; D43738; BAA07804.1; JOINED.

DR EMBL; D43739; BAA07804.1; JOINED.

DR EMBL; D43740; BAA07804.1; JOINED.

DR EMBL; D43741; BAA07804.1; JOINED.

DR EMBL; D43742; BAA07804.1; JOINED.

DR EMBL; D43743; BAA07804.1; JOINED.

DR EMBL; D43744; BAA07804.1; JOINED.

DR EMBL; D43745; BAA07804.1; JOINED.

DR EMBL; D43746; BAA07804.1; JOINED.

DR Genew; HGNC:1391; CACNA1D.

DR MIM; 114206; -

DR GO; GO:0005891; C:voltage-gated calcium channel complex; TAS.

DR GO; GO:0015270; F:dihydropyridine-sensitive calcium channel a. . .; TAS.

DR GO; GO:0006832; P:small molecule transport; TAS.

DR InterPro; IPR001682; Ca_Na_pore.

DR InterPro; IPR002077; Ca_channel.

DR InterPro; IPR002111; Cat_channel_TrpL.

DR InterPro; IPR005821; Ion_trans.

DR InterPro; IPR005446; LVDCCALPHAL.

DR InterPro; IPR005452; LVDCCALPHALD.

DR InterPro; IPR005820; M+channel_nlg.

DR Pfam; PF00520; ion_trans; 4.

DR PRINTS; PR00167; CACHANNEL.

DR PRINTS; PR01630; LVDCCALPHAL.

DR PRINTS; PR01636; LVDCCALPHALD.

DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;

DR Calcium channel; Glycoprotein; Repeat; Multigene family; Polymorphism;

DR Calcium-binding; Phosphorylation; Alternative splicing; Polymorphism;

DR Triplet repeat expansion.

FT REPEAT 113 409 I.

FT REPEAT 509 755 II.

FT REPEAT 873 1155 III.

FT REPEAT 1192 1467 IV.

FT DOMAIN 1 126 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 127 145 S1 OF REPEAT I (POTENTIAL).

FT DOMAIN 146 163 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 164 183 S2 OF REPEAT I (POTENTIAL).

FT DOMAIN 184 195 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 196 214 S3 OF REPEAT I (POTENTIAL).

FT DOMAIN 215 235 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 236 254 S4 OF REPEAT I (POTENTIAL).

FT DOMAIN 255 273 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 274 293 S5 OF REPEAT I (POTENTIAL).

FT DOMAIN 294 381 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 382 406 S6 OF REPEAT I (POTENTIAL).

FT	DOMAIN	407	523	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	524	543	S1 OF REPEAT II (POTENTIAL).
FT	DOMAIN	544	558	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	559	577	S2 OF REPEAT II (POTENTIAL).
FT	DOMAIN	578	585	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	586	604	S3 OF REPEAT II (POTENTIAL).
FT	DOMAIN	605	614	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	615	633	S4 OF REPEAT II (POTENTIAL).
FT	DOMAIN	634	652	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	653	673	S5 OF REPEAT II (POTENTIAL).
FT	DOMAIN	674	727	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	728	752	S6 OF REPEAT II (POTENTIAL).
FT	DOMAIN	753	886	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	887	905	S1 OF REPEAT III (POTENTIAL).
FT	DOMAIN	906	921	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	922	941	S2 OF REPEAT III (POTENTIAL).
FT	DOMAIN	942	953	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	954	972	S3 OF REPEAT III (POTENTIAL).
FT	DOMAIN	973	978	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	979	998	S4 OF REPEAT III (POTENTIAL).
FT	DOMAIN	999	1017	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1018	1037	S5 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1038	1127	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1128	1148	S6 OF REPEAT III (POTENTIAL).
FT	DOMAIN	1149	1205	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1206	1224	S1 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1225	1239	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1240	1259	S2 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1260	1266	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	1267	1288	S3 OF REPEAT IV (POTENTIAL).
FT	DOMAIN	1289	1313	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1314	1333	S4 OF REPEAT IV (POTENTIAL).
Query Match 79.4%; Score 27; DB 1; Length 2161;				
Best Local Similarity 83.3%; Pred. No. 82;				
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	MMMMR 6		
		:		
Db	3	MMMMK 8		
RESULT 5				
CCAD	RAT	STANDARD;	PRT; 2203 AA.	
ID	CCAD	RAT	STANDARD; Q09022; Q09023; Q09024; Q01542; Q62691; Q62815; Q63491;	
AC	P27732;	Q09022;	Q09023; Q09024; Q01542; Q62691; Q62815; Q63491;	
AC	Q63492;			
DT	01-OCT-1996	(Rel. 34, Created)		
DT	15-JUL-1999	(Rel. 38, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Voltage-dependent	L-type calcium channel alpha-1D subunit (Calcium		
DE	channel, L type, alpha-1 polypeptide, isoform 2) (RAT brain class D)	(RBD).		
DE	(RBD).			
GN	CACNA1D OR CACNL1A2 OR CCHL1A2 OR CACH3 OR CACN4.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RC	TISSUE=Insulinoma;			
RX	MEDLINE=95280950; PubMed=7760845;			
RA	Ihara Y., Yamada Y., Fujii Y., Gonoi T., Yano H., Yasuda K.,			
RA	Inagaki N., Seino Y., Seino S.;			
RT	"Molecular diversity and functional characterization of voltage-			
RT	dependent calcium channels (CACN4) expressed in pancreatic beta-			
RT	cells."			
RL	Mol. Endocrinol. 9:121-130(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS 3; 6; 7; 8 AND 14).			
RC	TISSUE=Brain;			
RX	MEDLINE=91299338; PubMed=1648940;			
RA	Hui A., Ellinor P.T., Krizanov O., Wang J.-J., Diebold R.J.,			

RA	Schwartz A.;			
RT	"Molecular cloning of multiple subtypes of a novel rat brain isoform			
RT	of the alpha-1 subunit of the voltage-dependent calcium channel."			
RL	Neuron 7:35-44(1991).			
RN	[3]			
RP	SEQUENCE OF 1-125 FROM N.A.			
RX	MEDLINE=96040125; PubMed=7553731;			
RA	Kamp T.J., Mitas M., Fields K.L., Asoh S., Chin H., Marban E.,			
RA	Nirenberg M.;			
RT	"Transcriptional regulation of the neuronal L-type calcium channel			
RT	alpha 1D subunit gene."			
RL	Cell. Mol. Neurobiol. 15:307-326(1995).			
RN	[4]			
RP	SEQUENCE OF 1100-1410 FROM N.A. (ISOFORMS 11 AND 12).			
RC	TISSUE=Kidney;			
RX	MEDLINE=93066265; PubMed=1279681;			
RA	Yu A.S.L., Hebert S.C., Brenner B.M., Lytton J.;			
RT	"Molecular characterization and nephron distribution of a family of			
RT	transcripts encoding the pore-forming subunit of Ca2+ channels in the			
RT	kidney."			
RL	Proc. Natl. Acad. Sci. U.S.A. 89:10494-10498(1992).			
RN	[5]			
RP	SEQUENCE OF 1218-1498 FROM N.A. (ISOFORM 13).			
RC	TISSUE=Osteosarcoma;			
RX	MEDLINE=96074617; PubMed=7479909;			
RA	Barry E.L.R., Gesek F.A., Froehner S.C., Friedman P.A.;			
RT	"Multiple calcium channel transcripts in rat osteosarcoma cells:			
RT	selective activation of alpha 1D isoform by parathyroid hormone."			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:10914-10918(1995).			
RN	[6]			
RP	SEQUENCE OF 1200-1493 FROM N.A. (ISOFORMS 3; 4; 5; 9 AND 10).			
RC	TISSUE=Hepatoma;			
RX	MEDLINE=97376179; PubMed=9232351;			
RA	Brereton H.M., Harland M.L., Frosio M., Petronijevic T.,			
RA	Barritt G.J.;			
RT	"Novel variants of voltage-operated calcium channel alpha-1 subunit			
RT	transcripts in a rat liver-derived cell line: deletion in the IVS4			
RT	voltage sensing region."			
RL	Cell Calcium 22:39-52(1997).			
RN	[7]			
RP	SEQUENCE OF 1307-1479 FROM N.A. (ISOFORM 3).			
RX	MEDLINE=90239020; PubMed=1692134;			
RA	Snutch T.P., Leonard J.P., Gilbert M.M., Lester H.A., Davidson N.;			
RT	"Rat brain expresses a heterogeneous family of calcium channels."			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:3391-3395(1990).			
CC	-1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE			
CC	ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED			
CC	IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE			
CC	CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,			
CC	CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D			
CC	GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)			
CC	CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)			
CC	GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),			
CC	PHENYLALKYLAMINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-III			
CC	(OMEGA-AGA-IIIa). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-			
CC	GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).			
CC	-1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT			
CC	COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS			
CC	IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-			
CC	FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS			
CC	SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM			
CC	CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA			
CC	LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=14;			
CC	Comment=The region sequenced in isoforms ROB3 and RKC5 is			
CC	identical to CACN4;			
CC	Name=1; Synonyms=CACN4A;			
CC	ISOId=P27732-1; Sequence=Displayed;			
CC	Name=2; Synonyms=CACN4B;			
CC	ISOId=P27732-2; Sequence=VSP_000923, VSP_000924;			
CC	Name=3; Synonyms=CACH3A, RB48, RBD-55;			

CC IsoId=P27732-3; Sequence=VSP_000921;
CC Name=4; Synonyms=Delta-IV-S3;
CC IsoId=P27732-4; Sequence=VSP_000919;
CC Name=5; Synonyms=Delta-IV-S4;
CC IsoId=P27732-5; Sequence=VSP_000922;
CC Name=6; Synonyms=RB9;
CC IsoId=P27732-6; Sequence=VSP_000920, VSP_000921;
CC Name=7; Synonyms=RB11;
CC IsoId=P27732-7; Sequence=VSP_000917;
CC Name=8; Synonyms=RB34;
CC IsoId=P27732-8; Sequence=VSP_000916;
CC Name=9; Synonyms=RH1;
CC IsoId=P27732-9; Sequence=VSP_000918;
CC Name=10; Synonyms=RH2;
CC IsoId=P27732-10; Sequence=VSP_000919, VSP_000922;
CC Name=11; Synonyms=RKC5;
CC IsoId=P27732-11; Sequence=Not described;
CC Name=12; Synonyms=RKC6;
CC IsoId=P27732-11; Sequence=VSP_000919;
CC Name=13; Synonyms=ROB3;
CC IsoId=P27732-14; Sequence=Not described;
CC Name=14; Synonyms=Truncated;
CC IsoId=P27732-12; Sequence=VSP_000925, VSP_000926;
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PANCREATIC ISLETS AND B-
CC LYMPHOCYTES.
CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
CC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; D38101; BAA07282.1; -
DR EMBL; D38102; BAA07283.1; -
DR EMBL; M57682; AAA42015.1; -
DR EMBL; U14005; AAB60515.1; -
DR EMBL; M99221; AAA40895.1; -
DR EMBL; U31772; AAA89156.1; -
DR EMBL; U49126; AAB61634.1; -
DR EMBL; U49127; AAB61635.1; -
DR EMBL; U49128; AAB61636.1; -
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002077; Ca_channel.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR003091; K_channel.
DR InterPro; IPR005446; LVDCCALPHAL.
DR InterPro; IPR005452; LVDCCALPHALD.
DR InterPro; IPR005820; M+channel_nlg.
DR Pfam; PF00520; ion_trans; 4.
DR PRINTS; PR00167; CCHANNEL.
DR PRINTS; PR00169; KCHANNEL.
DR PRINTS; PR01630; LVDCCALPHAL.
DR PRINTS; PR01636; LVDCCALPHALD.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel; Glycoprotein; Repeat; Multigene family;
KW Calcium-binding; Phosphorylation; Alternative splicing.
FT REPEAT 112 408 I.
FT REPEAT 528 774 II.
FT REPEAT 892 1174 III.
FT REPEAT 1211 1486 IV.
FT DOMAIN 1 126 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 127 145 S1 OF REPEAT I (POTENTIAL).
FT DOMAIN 146 163 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 164 183 S2 OF REPEAT I (POTENTIAL).
FT DOMAIN 184 195 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 196 214 S3 OF REPEAT I (POTENTIAL).
FT DOMAIN 215 235 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 236 254 S4 OF REPEAT I (POTENTIAL).
FT DOMAIN 255 273 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 274 293 S5 OF REPEAT I (POTENTIAL).
FT DOMAIN 294 381 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 382 406 S6 OF REPEAT I (POTENTIAL).
FT DOMAIN 407 582 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 583 602 S1 OF REPEAT II (POTENTIAL).
FT DOMAIN 603 617 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 618 636 S2 OF REPEAT II (POTENTIAL).
FT DOMAIN 637 644 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 645 663 S3 OF REPEAT II (POTENTIAL).
FT DOMAIN 664 673 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 674 692 S4 OF REPEAT II (POTENTIAL).
FT DOMAIN 693 711 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 712 732 S5 OF REPEAT II (POTENTIAL).
FT DOMAIN 733 786 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 787 811 S6 OF REPEAT II (POTENTIAL).
FT DOMAIN 812 945 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 946 964 S1 OF REPEAT III (POTENTIAL).
FT DOMAIN 965 980 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 981 1000 S2 OF REPEAT III (POTENTIAL).

Query Match 79.4%; Score 27; DB 1; Length 2203;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R 6
Db 3 M M M M M K 8

RESULT 6

RIF2_YEAST STANDARD; PRT; 395 AA.
ID RIF2_YEAST
AC Q06208;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE RIF2 protein (RAP1-interacting factor 2).
GN RIF2 OR YLR453C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Ansonge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierdeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
RN [2]
RP FUNCTION, AND SUBUNIT.
RX MEDLINE=97242430; PubMed=9087429;
RA Wotton D., Shore D.;
RT "A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to
RT regulate telomere length in Saccharomyces cerevisiae.";
RL Genes Dev. 11:748-760(1997).
CC -!- FUNCTION: Involved in transcriptional silencing and telomere

length regulation. Its role in telomere length regulation results from either a block in elongation or promoting degradation of the telomere ends. Loss of RIF1 function results in derepression of an HMR silencer, whose ARS consensus element has been deleted, and in the elongation of telomeres. RAP1 may target the binding of RIF1 to silencers and telomeres.

-!- SUBUNIT: Interacts with RIF1 and RAP1 C-terminus.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U22382; AAB67535.1; --
PIR; S55975; S55975.
SGD; S0004445; RIF2.
GO; GO:0005696; Citelomere; IDA.
GO; GO:0042162; F:telomeric DNA binding activity; IDA.
GO; GO:0007004; P:telomerase-dependent telomere maintenance; IMP.
Telomere.
SEQUENCE 395 AA; 45642 MW; FF6E3A22C8805DE9 CRC64;

Query Match 76.5%; Score 26; DB 1; Length 395;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
191 MMMVMR 196

RESULT 7
HRPN_ERWAM STANDARD; PRT; 403 AA.
ID HRPN_ERWAM
AC Q01099;
DT 01-APR-1993 (Rel. 25, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Harpin (Harpin-EA).
GN HRPN.
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC STRAIN=Ea321;
RX MEDLINE=92320301; PubMed=1621099;
RA Wei Z.-M., Laby R.J., Zumoff C.H., Bauer D.W., He S.Y., Collmer A., Beer S.V.;
RA "Harpin, elicitor of the hypersensitive response produced by the plant pathogen Erwinia amylovora.";
RL Science 257:85-88(1992).
RN [2]
RP REVISIONS.
RC STRAIN=Ea321;
RA Laby R.J., Kim J.F., Beer S.V.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: ELICITS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT UPON INFECTION. HARPIN ELICITS HR IN NON-HOSTS AND IS ALSO REQUIRED FOR PATHOGENICITY IN HOST PLANTS.
-!- SUBCELLULAR LOCATION: SECRETED; VIA THE HRP SECRETION PATHWAY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M92994; AAC31644.2; --
PIR; T08471; T08471.
KW Hypersensitive response.
FT DOMAIN 1 270 GLY-RICH.
FT DOMAIN 63 70 POLY-MET.
SQ SEQUENCE 403 AA; 39697 MW; 146FA642351D8E87 CRC64;

Query Match 76.5%; Score 26; DB 1; Length 403;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMMMRL 7
63 MMMMSM 69

RESULT 8
POPI_CAEEL STANDARD; PRT; 438 AA.
ID POPI_CAEEL
AC Q10666;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pop-1 protein.
GN POP-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=96069861; PubMed=7585963;
RA Lin R., Thompson S., Priess J.R.;
RT "pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos.";
RL Cell 83:599-609(1995).
RN [2]
RP REVISIONS.
RA Lin R., Thompson S., Priess J.R.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP FUNCTION, AND INTERACTION WITH POP-1.
RX PubMed=11742996;
RA Calvo D., Victor M., Gay F., Sui G., Luke M.P.-S., Dufourcq P., Wen G., Maduro M., Rothman J., Shi Y.;
RT "A POP-1 repressor complex restricts inappropriate cell type-specific gene transcription during Caenorhabditis elegans embryogenesis.";
RL EMBO J. 20:7197-7208(2001).
CC -!- FUNCTION: Essential for the specification of the mesodermal cell fate in early embryos. Represses expression of target genes via its interaction with hda-1 histone deacetylase.
CC -!- SUBUNIT: Interacts with hda-1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: Expressed maternally and zygotically.
CC -!- SIMILARITY: Contains 1 HMG box domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U37532; AAC05308.1; --
HSSP; P27782; 2LEF.
TRANSFAC; T03244; --
InterPro; IPR000910; HMG_12_box.
Pfam; PF00505; HMG_box; 1.
SMART; SM00398; HMG; 1.
PROSITE; PS50118; HMG_BOX_2; 1.

KW DNA-binding; Nuclear protein; Developmental protein.
 FT DOMAIN 55 173 PRO-RICH.
 FT DOMAIN 132 139 POLY-ALA.
 FT DNA BIND 192 262 HMG BOX.
 FT DOMAIN 340 347 POLY-SER.
 FT DOMAIN 359 362 POLY-GLN.
 FT DOMAIN 411 416 POLY-GLU.
 SQ SEQUENCE 438 AA; 48595 MW; F94073BE40B4095F CRC64;
 Query Match 76.5%; Score 26; DB 1; Length 438;
 Best Local Similarity 71.4%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 M M M M M R L 7
 Db 369 M L M Q M R L 375
 RESULT 9
 AMD2_XENLA
 ID AMD2_XENLA STANDARD; PRT; 875 AA.
 AC P12850;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Peptidyl-glycine alpha-amidating monooxygenase II precursor
 DE (EC 1.14.17.3) (Peptide C-terminal alpha-amidating enzyme II) (AE-II).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=88134244; PubMed=2829895;
 RA Ohshye K., Kitano K., Wada Y., Fuchimura K., Tanaka S., Mizuno K.,
 RA Matsuo H.;
 RT "Cloning of cDNA encoding a new peptide C-terminal alpha-amidating
 RT enzyme having a putative membrane-spanning domain from Xenopus laevis
 RT skin.";
 RL Biochem. Biophys. Res. Commun. 150:1275-1281(1988).
 CC -!- FUNCTION: C-terminal alpha-amidation of peptides. The reaction
 CC produces a peptidyl(2-hydroxyglycine) intermediate is unstable and
 CC dismutates to glyoxylate and the corresponding desglycine peptide
 CC amide.
 CC -!- CATALYTIC ACTIVITY: Peptidylglycine + ascorbate + O(2) =
 CC peptidyl(2-hydroxyglycine) + dehydroascorbate + H(2)O.
 CC -!- COFACTOR: COPPER AND ASCORBATE.
 CC -!- SIMILARITY: BELONGS TO THE COPPER TYPE II, ASCORBATE-DEPENDENT
 CC MONOOXYGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M19032; AAA49667.1; -.
 DR PIR; A27715; URXLA2.
 DR HSSP; P14925; 1PHM.
 DR InterPro; IPR001258; NHL.
 DR InterPro; IPR000323; Cu2_monooxygenase.
 DR InterPro; IPR000720; Pamonoxygenase.
 DR Pfam; PF03712; Cu2_monoox_C; 1.
 DR Pfam; PF01082; Cu2_monooxygen; 1.
 DR Pfam; PF01436; NHL; 4.
 DR PRINTS; PR00790; PAMONOXGNASE.
 DR PROSITE; PS00084; CU2_MONOOXYGENASE_1; 1.
 DR PROSITE; PS00085; CU2_MONOOXYGENASE_2; 1.
 KW Oxidoreductase; Monooxygenase; Copper; Vitamin C; Transmembrane;

KW Glycoprotein; Signal.
 FT SIGNAL 1 39 POTENTIAL.
 FT CHAIN 40 875 PEPTIDYL-GLYCINE ALPHA-AMIDATING
 FT MONOOXYGENASE II.
 FT INTRAGRANULAR (POTENTIAL).
 FT TRANSMEM 764 787 POTENTIAL.
 FT DOMAIN 40 763 POTENTIAL.
 FT DOMAIN 788 875 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 358 364 POLY-MET.
 FT DOMAIN 387 390 POLY-GLU.
 FT DOMAIN 852 856 POLY-ASP.
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 743 743 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 875 AA; 97084 MW; C07373AF6BF13450 CRC64;
 Query Match 76.5%; Score 26; DB 1; Length 875;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 M M M M M R L 7
 Db 358 M M M M M M M M 364
 RESULT 10
 NULM_RHISA
 ID NULM_RHISA STANDARD; PRT; 313 AA.
 AC O99824;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
 GN NDI.
 OS Rhipicephalus sanguineus (Brown dog tick).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
 OC NCBI_TaxID=34632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99083443; PubMed=9866211;
 RA Black W.C. IV, Roehrdanz R.L.;
 RT "Mitochondrial gene order is not conserved in arthropods: prostriate
 RT and metastriate tick mitochondrial genomes.";
 RL Mol. Biol. Evol. 15:1772-1785(1998).
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF081829; AAD05524.1; -.
 DR PIR; T11160; T11160.
 DR InterPro; IPR001694; Resp_NADH_dhl.
 DR Pfam; PF00146; NADHdh; 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 313 AA; 36721 MW; 846DA5BFB565EAD1 CRC64;
 Query Match 73.5%; Score 25; DB 1; Length 313;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 M M M M M M M 5
 Db 83 M M M M M M M 87

```
RESULT 11
AOX1_SOYBN STANDARD; PRT; 321 AA.
AC Q07185; Q41265;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alternative oxidase 1, mitochondrial precursor (EC 1.-.-.-).
GN AOX1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Shoot;
RX MEDLINE=94120028; PubMed=8290651;
RA Whelan J.M., McIntosh L., Day D.A.;
RT "Sequencing of a soybean alternative oxidase cDNA clone.";
RL Plant Physiol. 103:1481-1481(1993).
RN [2]
RP SEQUENCE OF 194-233 FROM N.A.
RX MEDLINE=96165778; PubMed=8580775;
RA Whelan J., Millar A.H., Day D.A.;
RT "The alternative oxidase is encoded in a multigene family in
soybean.";
RL Planta 198:197-201(1996).
CC -!- FUNCTION: CATALYZES CYANIDE-RESISTANT OXYGEN CONSUMPTION. MAY
CC INCREASE RESPIRATION WHEN THE CYTOCHROME RESPIRATORY PATHWAY IS
CC RESTRICTED, OR IN RESPONSE TO LOW TEMPERATURES (BY SIMILARITY).
CC -!- PATHWAY: Alternative respiratory pathway.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (PROBABLE).
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL, POSSIBLY IN THE INNER SURFACE
CC OF THE INNER MITOCHONDRIAL MEMBRANE.
CC -!- INDUCTION: By salicylic acid.
CC -!- SIMILARITY: BELONGS TO THE ALTERNATIVE OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X68702; CAA48653.1; -
DR EMBL; S81466; AAC34192.1; -
DR InterPro; IPR002680; AOX.
DR Pfam; PF01786; AOX; 1.
KW Oxidoreductase; Transit peptide; Mitochondrion; Respiratory chain;
KW Inner membrane; Transmembrane; Multigene family.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? ? ALTERNATIVE OXIDASE 1.
FT CHAIN ? 321 ?
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT DISULFID 94 94 INTERCHAIN (POTENTIAL).
FT METAL 112 112 POTENTIAL.
FT METAL 165 165 POTENTIAL.
FT METAL 192 192 POTENTIAL.
FT METAL 233 233 POTENTIAL.
FT METAL 294 294 POTENTIAL.
FT METAL 299 299 POTENTIAL.
FT DOMAIN 1 5 POLY-MET.
SQ SEQUENCE 321 AA; 36437 MW; EEOFF981F0C2D0E7 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 321;
Best Local Similarity 100.0%; Pred.No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M 5
|||||
```

```
Db 1 M M M M M 5

RESULT 12
CY1_HUMAN STANDARD; PRT; 325 AA.
ID CY1_HUMAN
AC P08574;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cytochrome c1, heme protein, mitochondrial precursor (Cytochrome c-1).
GN CYC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89109139; PubMed=2536365;
RA Suzuki H., Hosokawa Y., Nishikimi M., Ozawa T.;
RT "Structural organization of the human mitochondrial cytochrome c1
gene.";
RL J. Biol. Chem. 264:1368-1374(1989).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT VAL-89.
RX MEDLINE=88233946; PubMed=2836796;
RA Nishikimi M., Ohta S., Suzuki H., Tanaka T., Kikkawa F.,
RA Tanaka M., Kagawa Y., Ozawa T.;
RT "Nucleotide sequence of a cDNA encoding the precursor to human
cytochrome c1.";
RL Nucleic Acids Res. 16:3577-3577(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, Lung, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 99-325 FROM N.A.
RX MEDLINE=87241521; PubMed=3036122;
RA Nishikimi M., Suzuki H., Ohta S., Sakurai T., Shimomura Y.,
RA Tanaka M., Kagawa Y., Ozawa T.;
RT "Isolation of a cDNA clone for human cytochrome c1 from a lambda gt11
expression library.";
RL Biochem. Biophys. Res. Commun. 145:34-39(1987).
RN [5]
RP VARIANT VAL-89.
RX MEDLINE=99381536; PubMed=10453733;
RA Valnot I., Kassis J., Chretien D., de Lonlay P., Parfait B.,
RA Munnich A., Kachaner J., Rustin P., Roetig A.;
RT "A mitochondrial cytochrome b mutation but no mutations of nuclearly
encoded subunits in ubiquinol cytochrome c reductase (complex III)
deficiency.";
RL Hum. Genet. 104:460-466(1999).
CC -!- FUNCTION: THIS IS THE HEME-CONTAINING COMPONENT OF THE CYTOCHROME
```


CC B-C1 COMPLEX, WHICH ACCEPTS ELECTRONS FROM RIESKE PROTEIN AND
 CC TRANSFERS ELECTRONS TO CYTOCHROME C IN THE MITOCHONDRIAL
 CC RESPIRATORY CHAIN.
 CC -!- SUBUNIT: BCL COMPLEX CONTAINS 11 SUBUNITS; 3 RESPIRATORY SUBUNITS
 CC (CYTOCHROME B, CYTOCHROME C1, RIESKE PROTEIN), 2 CORE PROTEINS AND
 CC 6 LOW-MOLECULAR WEIGHT PROTEINS.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial intermembrane space.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J04444; AAA52135.1; -
 CC EMBL; M16597; AAA35730.1; -
 CC EMBL; BC001006; AAH01006.1; -
 CC EMBL; BC015616; AAH15616.1; -
 CC EMBL; BC020566; AAH20566.1; -
 CC EMBL; X06994; CAA30052.1; -
 CC PIR; A31481; S00680.
 CC Genew; HGNC:2579; CYC1.
 CC MIM; 123980; -
 CC GO; GO:0005739; C:Mitochondrion; TAS.
 CC GO; GO:0009461; F:Cytochrome c; TAS.
 CC InterPro; IPR002326; Cyt C1.
 CC InterPro; IPR000345; CytC_heme_bind.
 CC Pfam; PF02167; Cytochrome C1; 1.
 CC PRINTS; PR00603; CYTOCHROME1.
 CC PROSITE; PS00190; CYTOCHROME C; 1.
 CC Electron transport; Respiratory chain; Heme; Mitochondrion;
 CC Transmembrane; Transit peptide; Polymorphism.
 CC TRANSIT 1 84 MITOCHONDRION.
 CC CHAIN 85 325 CYTOCHROME C1, HEME PROTEIN.
 CC BINDING 121 121 HEME (COVALENT).
 CC BINDING 124 124 HEME (COVALENT).
 CC METAL 125 125 IRON (HEME AXIAL LIGAND).
 CC METAL 244 244 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC TRANSMEM 292 306 ANCHORS TO THE MEMBRANE (POTENTIAL).
 CC VARIANT 89 89 L -> V.
 CC
 CC SEQUENCE 325 AA; 35390 MW; CC8815E60E99EBDC CRC64;
 CC
 CC Query Match 73.5%; Score 25; DB 1; Length 325;
 CC Best Local Similarity 71.4%; Pred. No. 43;
 CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 MMTMMRL 7
 CC Db 292 MLMMML 298
 CC
 CC RESULT 13
 CC HKL6_LYCES
 CC ID HKL6_LYCES STANDARD; PRT; 355 AA.
 CC AC O22299;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Homeobox protein knotted-1 like LET6.
 CC GN LET6.
 CC OS Lycopersicon esculentum (Tomato).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 CC OX NCBI_TaxID=4081;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=CV. VENT Cherry;
 CC RX MEDLINE=98145476; PubMed=9484482;
 CC RA Janssen B.J., Williams A., Chen J.J., Mathern J., Hake S., Sinha N.;

RT "Isolation and characterization of two knotted-like homeobox genes
 RT from tomato.";
 RT Plant Mol. Biol. 36:417-425(1998).
 CC -!- FUNCTION: MAY HAVE A ROLE TO PLAY IN FORMATIVE EVENTS IN OVULE AND
 CC EMBRYO MORPHOGENESIS. PROBABLY BINDS TO THE DNA SEQUENCE 5'-TGAC-
 CC 3'.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING LATERAL ORGANS AND
 CC DEVELOPING OVARIES IN FLOWERS.
 CC -!- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEBOX FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF000141; AAC49917.1; -
 CC PIR; T04317; T04317.
 CC HSSP; P40424; 1B72.
 CC TRANSFAC; T04057; -
 CC InterPro; IPR005539; ELK.
 CC InterPro; IPR001356; Homeobox.
 CC InterPro; IPR005540; KNOX1.
 CC InterPro; IPR005541; KNOX2.
 CC Pfam; PF03789; ELK; 1.
 CC Pfam; PF00046; homeobox; 1.
 CC Pfam; PF03790; KNOX1; 1.
 CC Pfam; PF03791; KNOX2; 1.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC KW Homeobox; DNA-binding; Nuclear protein.
 CC FT DOMAIN 24 33 POLY-ASN.
 CC FT DOMAIN 47 51 POLY-MET.
 CC FT DOMAIN 52 56 POLY-PRO.
 CC FT DOMAIN 60 63 POLY-ASN.
 CC FT DOMAIN 79 84 POLY-ASN.
 CC FT DOMAIN 91 95 POLY-SER.
 CC FT DOMAIN 142 146 POLY-SER.
 CC FT DOMAIN 235 258 ELK DOMAIN.
 CC FT DNA BIND 259 321 HOMEBOX (TALE-TYPE).
 CC SQ SEQUENCE 355 AA; 39796 MW; 9E9A1AFD75808C49 CRC64;
 CC
 CC Query Match 73.5%; Score 25; DB 1; Length 355;
 CC Best Local Similarity 100.0%; Pred. No. 47;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 MMTMM 5
 CC Db 47 MMTMM 51
 CC
 CC RESULT 14
 CC Y4WD_RHISN
 CC ID Y4WD_RHISN STANDARD; PRT; 377 AA.
 CC AC P55682;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 01-NOV-1997 (Rel. 35, Last annotation update)
 CC DE Hypothetical transport protein Y4WD.
 CC GN Y4WD.
 CC OS Rhizobium sp. (strain NGR234).
 CC OG Plasmid sym pNGR234a.
 CC OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
 CC OX NCBI_TaxID=394;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=97305956; PubMed=9163424;

RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- FUNCTION: COULD BE INVOLVED IN A TRANSPORT SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: TO R.MELILOTI MOSC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB000103; AAB91911.1; --
DR InterPro; IPR007114; MFS.
KW Hypothetical protein; Transmembrane; Transport; Plasmid.
FT TRANSMEM 4 24 POTENTIAL.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 134 154 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 278 298 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 356 376 POTENTIAL.
SQ SEQUENCE 377 AA; 39051 MW; 49CF6E44AA0D74BD CRC64;

Query Match 73.5%; Score 25; DB 1; Length 377;
Best Local Similarity 83.3%; Pred. No. 50;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MVMRL 7
Db 250 MVMRL 255
|:|||||

RESULT 15
STM ARATH
ID STM ARATH STANDARD; PRT; 382 AA.
AC Q38874; Q8RXJ1; Q9MAV3;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homeobox protein SHOOT MERISTEMLESS.
GN STM OR ATIG52360 OR F2401.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wassilewskija;
RX MEDLINE=96135134; PubMed=8538741;
RA Long J.A., Moan E.I., Medford J.I., Barton M.K.;
RT "A member of the KNOTTED class of homeodomain proteins encoded by the
RT STM gene of Arabidopsis.";
RL Nature 379:66-69(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.X.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,

RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
RN [3]
RP SEQUENCE OF 57-382 FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION.
RX MEDLINE=21932426; PubMed=11934861;
RA Byrne M.E., Simorowski J., Martienssen R.A.;
RT "ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis.";
RL Development 129:1957-1965(2002).
CC -!- FUNCTION: Required for shoot apical meristem formation during
CC embryogenesis. Negatively regulates ASYMMETRIC LEAVES1 (AS1) and
CC ASYMMETRIC LEAVES2 (AS2 or LBD6). Probably binds to the DNA
CC sequence 5'-TGAC-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in all four types of shoot apical
CC meristems (SAM) i.e., in vegetative, axillary, inflorescence and
CC floral.
CC -!- DEVELOPMENTAL STAGE: First expressed in early to mid-globular-
CC stage embryos. In late globular stage, detected as a stripe
CC running medially across the top of the embryo. In heart stage
CC embryo, expression is restricted to a notch between the
CC cotyledons. In seedlings and adult plants found in all shoot
CC apical meristems. In the inflorescence meristem, expression
CC disappears as floral buds are initiated and reappears in the later
CC floral meristem where it is found in the central portion of the
CC developing gynoecium. Also detected in the L1 layer of embryo.
CC -!- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32344; AAC49148.1; --
DR EMBL; AC003113; AAF70849.1; --
DR EMBL; AY080857; AAL87330.1; --
DR PIR; S68456; S68456.
DR PIR; T01446; T01446.
DR HSSP; P41778; 1DU6.
DR InterPro; IPR005539; ELK.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR005540; KNOX1.
DR InterPro; IPR005541; KNOX2.
DR Pfam; PF03789; ELK; 1.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein; Developmental protein.
FT DOMAIN 260 283 ELK DOMAIN.
FT DNA_BIND 284 346 HOMEBOX (TALE-TYPE).
FT

FT	DOMAIN	26	30	POLY-MET.
FT	DOMAIN	37	60	HIS-RICH.
FT	DOMAIN	83	89	POLY-SER.
FT	DOMAIN	113	120	POLY-SER.
FT	DOMAIN	161	166	POLY-ALA.
FT	CONFLICT	115	115	S -> F (IN REF. 1).
FT	CONFLICT	229	233	MISSING (IN REF. 2).
FT	CONFLICT	369	369	G -> D (IN REF. 1).
SQ	SEQUENCE	382 AA;	42753 MW;	6227D3DE4093E732 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MMMMM	5
Db	26	MMMMM	30

Search completed: February 11, 2004, 17:04:17
Job time : 6.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-12

Perfect score: 34

Sequence: 1 MWMVMRL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.**

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	70	T48381	hypothetical prote
2	30	88.2	192	H96828	hypothetical prote
3	30	88.2	249	T08410	hypothetical prote
4	30	88.2	353	T07725	hypothetical prote
5	27	79.4	39	PC4294	high mobility grou
6	27	79.4	146	S14946	2S seed storage pr
7	27	79.4	147	JQ1320	high mobility grou
8	27	79.4	427	T04869	transforming prote
9	27	79.4	1610	A46227	voltage-dependent
10	27	79.4	1646	JH0432	voltage-dependent
11	27	79.4	2161	JH0564	calcium channel al
12	27	79.4	2181	A38198	calcium channel al
13	27	79.4	2203	T42742	voltage-dependent
14	26	76.5	216	H82635	DNA-3-methyladenin
15	26	76.5	302	H96811	protein F3F9.20 [i
16	26	76.5	309	E83999	mutants block spor
17	26	76.5	340	C87732	protein W03D8.4 [i
18	26	76.5	340	T32931	hypothetical prote
19	26	76.5	375	T06096	hypothetical prote
20	26	76.5	395	S55975	hypothetical prote
21	26	76.5	403	T08471	harpin - Erwinia a
22	26	76.5	436	E96635	hypothetical prote
23	26	76.5	438	A57667	pop-1 protein - Ca
24	26	76.5	484	H95377	probable TrkH-like
25	26	76.5	875	URXLA2	peptidylglycine mo
26	25	73.5	68	T03645	hypothetical prote
27	25	73.5	98	A98222	hypothetical prote
28	25	73.5	120	AI2181	hypothetical prote
29	25	73.5	124	F81942	probable regulator

30	25	73.5	124	2	A81163	conserved hypothet
31	25	73.5	135	2	F86239	protein F20B24.4 [
32	25	73.5	176	2	A86159	hypothetical prote
33	25	73.5	197	2	T26156	hypothetical prote
34	25	73.5	217	2	H86657	conserved hypothet
35	25	73.5	220	2	AD0225	probable amino-aci
36	25	73.5	242	2	B82815	conserved hypothet
37	25	73.5	268	2	T04787	hypothetical prote
38	25	73.5	275	2	T51651	myb-related transc
39	25	73.5	289	2	T29216	hypothetical prote
40	25	73.5	292	2	T09030	hypothetical prote
41	25	73.5	307	2	AD0883	conserved hypothet
42	25	73.5	313	2	T11160	NADH2 dehydrogenas
43	25	73.5	321	2	S31711	alternative respir
44	25	73.5	325	1	S00680	ubiquinol-cytochro
45	25	73.5	327	2	T09687	chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

T48381
hypothetical protein F12E4.330 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48381
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24492
A:Accession: T48381
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-70 <BEV>
A:Cross-references: EMBL:AL162751
A:Experimental source: cultivar Columbia; BAC clone F12E4
C:Genetics:
A:Map position: 5
A:Note: F12E4.330

Query Match 94.1%; Score 32; DB 2; Length 70;
Best Local Similarity 85.7%; Pred. No. 0.57;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MWMVMRL 7
DB	3	MWMVMRI 9

RESULT 2

H96828
hypothetical protein F19K16.24 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 21-Oct-2002
C:Accession: H96828
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H96828
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <STO>
A:Cross-references: GB:AE005173; NID:96453865; PIDN:AAF09049.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19K16.24

A;Map position: 1
C;Superfamily: stellacyanin

Query Match 88.2%; Score 30; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 2 M M M M M M R 7

RESULT 3
T08410
hypothetical protein F18B3.170 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C;Accession: T08410
R;Querier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salanoub
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16409
A;Accession: T08410
A;Molecule type: DNA
A;Residues: 1-249 <QUE>
A;Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.170
A;Experimental source: cultivar Columbia; BAC clone F18B3
C;Genetics:
A;Gene: ATSP:F18B3.170
A;Map position: 3

Query Match 88.2%; Score 30; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 174 M M M M M M R 179

RESULT 4
T07725
hypothetical protein T23J7.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 22-Oct-1999
C;Accession: T07725
R;Querier, F.; Choisine, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigu
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15793
A;Accession: T07725
A;Molecule type: DNA
A;Residues: 1-353 <QUE>
A;Cross-references: EMBL:AL049746; GSPDB:GN00061; ATSP:T23J7.190
A;Experimental source: cultivar Columbia; BAC clone T23J7
C;Genetics:
A;Gene: ATSP:T23J7.190
A;Map position: 3
A;Introns: 28/2; 83/3; 165/3; 262/3

Query Match 88.2%; Score 30; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 122 M M M M M M R 127

RESULT 5
PC4294
high mobility group protein pf16 - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jun-2000
C;Accession: PC4294

R;Nambiar, A.; Kandala, J.C.; Dolan, S.A.; Jensen, J.B.; Guntaka, R.V.
Biochem. Biophys. Res. Commun. 234, 101-106, 1997
A;Title: Molecular cloning and characterization of a cDNA for the highly conserved HMG-
A;Reference number: JCS403; MUID:97312528; PMID:9168969
A;Accession: PC4294
A;Molecule type: mRNA
A;Residues: 1-39 <NAM>
C;Comment: This protein is involved in DNA replication, chromatin assembly and transcri
C;Genetics:
A;Gene: Pf16

Query Match 79.4%; Score 27; DB 2; Length 39;
Best Local Similarity 83.3%; Pred. No. 5.2;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 21 M M M M M M K 26

RESULT 6
S14946
2S seed storage protein large chain - Brazil nut
N;Alternate names: albumin 2S precursor
C;Species: Bertholletia excelsa (Brazil nut)
C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C;Accession: S14946; S14479; S06252; S21640; B25802
R;Gander, E.S.; Holmstroem, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Gros
Plant Mol. Biol. 16, 437-448, 1991
A;Title: Isolation, characterization and expression of a gene coding for a 2S albumin f
A;Reference number: S14946; MUID:91370890; PMID:1840683
A;Accession: S14946
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <GAN>
A;Cross-references: EMBL:X54490; NID:gl7710; PIDN:CAA38362.1; PID:gl7711
A;Note: the authors translated the codon CTT for residue 13 as Val and GTC for residue
R;Basuener, R.; Schlesier, B.
submitted to the EMBL Data Library, December 1990
A;Reference number: S14479
A;Accession: S14479
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <BAS>
A;Cross-references: EMBL:X57027; NID:gl7714; PID:gl7715
R;Altenbach, S.B.; Pearson, K.W.; Leung, F.W.; Sun, S.S.M.
Plant Mol. Biol. 8, 239-250, 1987
A;Title: Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptic
A;Reference number: S06252
A;Accession: S06252
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-146 <ALT>
A;Experimental source: clone pHS-3
A;Note: part of this sequence, including the amino end of the large chain, was confirme
R;Basuener, R.; Schlesier, B.
submitted to the EMBL Data Library, December 1990
A;Reference number: S21640
A;Accession: S21640
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <BA2>
A;Cross-references: EMBL:X57028; NID:gl7716; PID:gl7717
R;Ampe, C.; Van Damme, J.; de Castro, L.A.B.; Sampaio, M.J.A.M.; Van Montagu, M.; Vande
Eur. J. Biochem. 159, 597-604, 1986
A;Title: The amino-acid sequence of the 2S sulphur-rich proteins from seeds of Brazil r
A;Reference number: A91173; MUID:87004679; PMID:3758080
A;Accession: B25802
A;Molecule type: Protein
A;Residues: 70-90, 'E', 92-121, 'M', 123-125, 'L', 127-142 <AMP>
C;Genetics:
A;Introns: 60/3
C;Superfamily: wheat alpha-amylase inhibitor

F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-46/Domain: propeptide #status predicted <PRO>
F;47-69/Product: seed storage protein small chain #status predicted <SCH>
F;70-146/Product: seed storage protein large chain #status experimental <LCH>

Query Match 79.4%; Score 27; DB 2; Length 146;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MMMMRRL 7
Db 99 MMMMRM 104

RESULT 7

JQ1320
high mobility group protein Pf16 - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jun-2000
C;Accession: JQ1320

R;Guntaka, R.V.; Kandala, J.C.; Reddy, V.D.
Biochem. Biophys. Res. Commun. 182, 412-419, 1992
A;Title: Cloning and characterization of a highly conserved HMG-like protein (PF16) gene
A;Reference number: JQ1320; MUID:92118046; PMID:1731798

A;Accession: JQ1320
A;Molecule type: DNA
A;Residues: 1-147 <GUN>
A;Cross-references: GB:M86518; NID:gl60325; PID:gl60326
A;Experimental source: strain FCR/3

C;Comment: This protein interacts with other nuclear proteins and serve as a transcription factor
C;Genetics:

A;Gene: pf16
F;7-18,19-30/Region: duplication
F;70-91/Region: aspartic acid/glutamic acid-rich
F;126-133/Region: basic

Query Match 79.4%; Score 27; DB 2; Length 147;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMR 6
Db 11 MMMMK 16

RESULT 8

T04869
transforming protein myb homolog F28A21.180 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 19-May-2000
C;Accession: T04869

R;Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer, R.
Submitted to the Protein Sequence Database, February 1999
A;Reference number: 215387

A;Accession: T04869
A;Molecule type: DNA
A;Residues: 1-427 <BEV>

A;Cross-references: EMBL:AL035526
A;Experimental source: cultivar Columbia; BAC clone F28A21
C;Genetics:

A;Map position: 4
A;Introns: 226/2; 268/3
A;Note: F28A21.180

C;Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolog F;212-263/Domain: myb DNA-binding repeat homology <MYB>

Query Match 79.4%; Score 27; DB 2; Length 427;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MMMMR 7
Db 145 MMMMRM 150

RESULT 9

A46227
voltage-dependent Ca2+ channel alpha 1-subunit - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 27-Oct-1993 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000
C;Accession: A46227

R;Yaney, G.C.; Wheeler, M.B.; Wei, X.; Perez-Reyes, E.; Birnbaumer, L.; Boyd III, A.E.; Mol. Endocrinol. 6, 2143-2152, 1992

A;Title: Cloning of a novel alpha 1-subunit of the voltage-dependent calcium channel f
A;Reference number: A46227; MUID:93149124; PMID:1337146

A;Accession: A46227

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1610 <YAN>

A;Experimental source: insulin-secreting cell line HIT-T15

A;Note: sequence extracted from NCBI backbone (NCBI:123692)

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 79.4%; Score 27; DB 2; Length 1610;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMR 6
Db 2 MMMMK 7

RESULT 10

JH0422

voltage-dependent calcium channel complex alpha-1 chain - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jan-2000

C;Accession: JH0422; D35901; I60901

R;Hui, A.; Ellinor, P.T.; Krizanov, O.; Wang, J.J.; Diebold, R.J.; Schwartz, A.

Neuron 7, 35-44, 1991

A;Title: Molecular cloning of multiple subtypes of a novel rat brain isoform of the al

A;Reference number: JH0422; MUID:91299338; PMID:1648940

A;Accession: JH0422

A;Molecule type: mRNA

A;Residues: 1-1646 <HUI>

A;Cross-references: GB:M57682; NID:G206573; PIDN:AAA42015.1; PID:G206574

A;Experimental source: brain

R;Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.; Davidson, N.

Proc. Natl. Acad. Sci. U.S.A. 87, 3391-3395, 1990

A;Title: Rat brain expresses a heterogeneous family of calcium channels.

A;Reference number: A35901; MUID:90239020; PMID:1692134

A;Accession: D35901

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A;Molecule type: mRNA

A;Residues: 1247-1434 <SNU>

A;Experimental source: brain

R;Yu, A.S.L.; Hebert, S.C.; Brenner, B.M.; Lytton, J.

Proc. Natl. Acad. Sci. U.S.A. 89, 10494-10498, 1992

A;Title: Molecular characterization and nephron distribution of a family of transcripts

A;Reference number: A46422; MUID:93066265; PMID:1279681

A;Accession: I60901

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1040-1261,1305-1365 <RES>

A;Cross-references: GB:M99221; NID:G203370; PIDN:AAA40895.1; PID:G203371

A;Experimental source: kidney

C;Comment: Calcium channels are essential for many cellular functions, such as muscle c

n.

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

C;Keywords: alternative splicing; calcium binding; calcium channel; glycoprotein; phosph

F;1463-1491/Domain: calcium binding #status predicted <EFC>

F;154,224,328/Binding site: carboxylate (Asn) (covalent) #status predicted

F;464,848,1489,1584/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 79.4%; Score 27; DB 2; Length 1646;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
| | | | |
Db 2 MMMMMK 7

RESULT 11

JH0564

calcium channel alpha-1D chain - human

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 24-Nov-1999

C;Accession: JH0564

R;Williams, M.E.; Feldman, D.H.; McCue, A.F.; Brenner, R.; Velicelebi, G.; Ellis, S.B.; Neuron 8, 71-84, 1992

A;Title: Structure and functional expression of alpha1, alpha2, and beta subunits of a neuron calcium channel

A;Reference number: JH0564; MUID:92110010; PMID:1309651

A;Accession: JH0564

A;Molecule type: mRNA

A;Residues: 1-2161 <WIL>

A;Cross-references: GB:M76558

A;Experimental source: GB:M76558

C;Comment: This protein is a subunit of the voltage-dependent calcium channel.

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

C;Keywords: transmembrane protein

F;127-145/Domain: transmembrane #status predicted <IS1>

F;164-183/Domain: transmembrane #status predicted <IS2>

F;195-214/Domain: transmembrane #status predicted <IS3>

F;236-254/Domain: transmembrane #status predicted <IS4>

F;274-293/Domain: transmembrane #status predicted <IS5>

F;382-406/Domain: transmembrane #status predicted <IS6>

F;524-542/Domain: transmembrane #status predicted <II1>

F;558-577/Domain: transmembrane #status predicted <II2>

F;586-603/Domain: transmembrane #status predicted <II3>

F;615-633/Domain: transmembrane #status predicted <II4>

F;653-672/Domain: transmembrane #status predicted <II5>

F;728-752/Domain: transmembrane #status predicted <II6>

F;888-905/Domain: transmembrane #status predicted <S11>

F;922-941/Domain: transmembrane #status predicted <S21>

F;954-972/Domain: transmembrane #status predicted <S31>

F;980-998/Domain: transmembrane #status predicted <S41>

F;1018-1037/Domain: transmembrane #status predicted <S51>

F;1128-1152/Domain: transmembrane #status predicted <S61>

F;1206-1224/Domain: transmembrane #status predicted <VS1>

F;1240-1259/Domain: transmembrane #status predicted <VS2>

F;1268-1286/Domain: transmembrane #status predicted <VS3>

F;1315-1333/Domain: transmembrane #status predicted <VS4>

F;1353-1372/Domain: transmembrane #status predicted <VS5>

F;1440-1464/Domain: transmembrane #status predicted <VS6>

Query Match 79.4%; Score 27; DB 2; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
| | | | |
Db 3 MMMMMK 8

RESULT 12

A38198

calcium channel alpha-1 chain, pancreatic - human

N;Alternate names: beta-cell-type calcium channel alpha-1 chain; neuroendocrine-type calcium channel

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000

C;Accession: A38198

R;Seino, S.; Chen, L.; Seino, M.; Blondel, O.; Takeda, J.; Johnson, J.H.; Bell, G.I. Proc. Natl. Acad. Sci. U.S.A. 89, 584-588, 1992

A;Title: Cloning of the alpha1 subunit of a voltage-dependent calcium channel expressed in oocytes

A;Reference number: A38198; MUID:92115705; PMID:1309948

A;Accession: A38198

A;Molecule type: mRNA

A;Residues: 1-2181 <SEI>

A;Cross-references: GB:M83566; NID:g179751; PIDN:AAA35629.1; PID:g179752

A;Experimental source: pancreatic beta cells

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

C;Keywords: membrane protein; voltage-gated ion channel

Query Match 79.4%; Score 27; DB 2; Length 2181;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
| | | | |
Db 3 MMMMMK 8

RESULT 13

T42742

voltage-dependent calcium channel alpha 1 chain, isoform CACN4A - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C;Accession: T42742

R;Ihara, Y.; Yamada, Y.; Fujii, Y.; Gonori, T.; Yano, H.; Yasuda, K.; Inagaki, N.; Seir Mol. Endocrinol. 9, 121-130, 1995

A;Title: Molecular diversity and functional characterization of voltage-dependent calcium channels in rat brain

A;Reference number: Z22258; MUID:95280950; PMID:7760845

A;Accession: T42742

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-2203 <IHA>

A;Cross-references: EMBL:D38101; NID:g736711; PIDN:BAA07282.1; PID:g736712

A;Experimental source: insulinoma RINm5F complementary DNA library

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

C;Keywords: calcium channel

Query Match 79.4%; Score 27; DB 2; Length 2203;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
| | | | |
Db 3 MMMMMK 8

RESULT 14

H82695

DNA-3-methyladenine glycosidase XF1326 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C;Accession: H82695

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: H82695

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-216 <SIM>

A;Cross-references: GB:AE003965; GB:AE003849; NID:g9106313; PIDN:AAF84135.1; GSPDB:GN000000000

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laiz Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A.; Authors: Martins, E.M.P.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawas A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; A;Reference number: A59328

A;Contents: annotation

C;Genetics:
A;Gene: XF1326

Query Match 76.5%; Score 26; DB 2; Length 216;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 1 MMTMMRL 7
Db 139 MMLMRL 145

RESULT 15
H96811
protein F3F9.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: H96811
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96811
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <STO>
A;Cross-references: GB:AE005173; NID:G8052542; PIDN:AAF71806.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3F9.20
A;Map position: 1

Query Match 76.5%; Score 26; DB 2; Length 302;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
Db 159 MMTMMI 165

Search completed: February 11, 2004, 17:11:56
Job time : 11.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 MWMWRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27	79.4	10	1	US-08-417-279-2
2	27	79.4	10	5	PCT-US94-14173-2
3	27	79.4	61	4	US-09-252-991A-29802
4	27	79.4	2161	1	US-07-745-206A-2
5	27	79.4	2161	1	US-08-455-543A-49
6	27	79.4	2161	1	US-08-455-543A-51
7	27	79.4	2161	2	US-08-223-305C-49
8	27	79.4	2161	2	US-08-223-305C-51
9	27	79.4	2161	2	US-08-311-363-2
10	26	76.5	385	1	US-08-891-254-3
11	26	76.5	385	2	US-08-819-539-3
12	26	76.5	385	5	PCT-US93-06243-2
13	26	76.5	385	5	PCT-US96-08819-3
14	26	76.5	403	2	US-08-200-724A-2
15	26	76.5	403	2	US-09-030-270A-3
16	26	76.5	403	3	US-08-851-376A-2
17	26	76.5	403	3	US-08-984-207-3
18	26	76.5	403	3	US-09-013-587-3
19	26	76.5	403	4	US-09-086-118-23
20	26	76.5	989	2	US-08-070-301-16
21	25	73.5	62	1	US-08-488-961-6
22	25	73.5	62	3	US-08-973-297-6
23	25	73.5	62	5	PCT-US96-06511-6
24	25	73.5	67	3	US-09-120-365-98
25	25	73.5	67	3	US-09-515-039-98
26	25	73.5	185	4	US-09-328-352-6508
27	25	73.5	284	4	US-09-328-352-4466

28	25	73.5	521	4	US-09-107-532A-3961	Sequence 3961, Ap
29	25	73.5	935	1	US-07-707-367-2	Sequence 2, Appli
30	25	73.5	1160	4	US-09-328-352-6457	Sequence 6457, Ap
31	24	70.6	110	4	US-09-252-991A-16573	Sequence 16573, A
32	24	70.6	190	4	US-09-328-352-4549	Sequence 4549, Ap
33	24	70.6	299	4	US-09-252-991A-22835	Sequence 22835, A
34	24	70.6	312	4	US-09-107-532A-6637	Sequence 6637, Ap
35	24	70.6	393	4	US-09-252-991A-16771	Sequence 16771, A
36	24	70.6	429	4	US-09-252-991A-16841	Sequence 16841, A
37	24	70.6	490	4	US-09-252-991A-29975	Sequence 29975, A
38	24	70.6	668	4	US-09-252-991A-22341	Sequence 22341, A
39	24	70.6	687	2	US-08-540-804-4	Sequence 4, Appli
40	24	70.6	687	2	US-08-218-265-4	Sequence 4, Appli
41	24	70.6	687	3	US-08-521-872-4	Sequence 4, Appli
42	24	70.6	687	3	US-08-590-399-4	Sequence 4, Appli
43	23	67.6	88	4	US-09-328-352-5555	Sequence 5555, Ap
44	23	67.6	93	4	US-09-134-001C-5186	Sequence 5186, Ap
45	23	67.6	189	3	US-09-105-343A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-417-279-2
; Sequence 2, Application US/08417279
; Patent No. 5543498
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; APPLICANT: Igarashi, Michihiro
; TITLE OF INVENTION: Peptides to Overcome Inhibition of Nerve
; TITLE OF INVENTION: Growth
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,279
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/166,350
; FILING DATE: 14-DEC-1993
; APPLICATION NUMBER: 08/162,480
; FILING DATE: 07-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3960001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: both

Query Match 79.4%; Score 27; DB 1; Length 10;
Best Local Similarity 83.3%; Pred. No. 0.94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWMWRL 6

Db 1 MLMMMR 6

RESULT 2
PCT-US94-14173-2
; Sequence 2, Application PC/TUS9414173
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Peptides to Overcome Inhibition of Nerve
; TITLE OF INVENTION: Growth
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14173
; FILING DATE: 06-DEC-1994
; PRIOR APPLICATION NUMBER: US 08/166,350
; FILING DATE: 14-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/162,480
; FILING DATE: 07-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3960001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; PCT-US94-14173-2

Query Match 79.4%; Score 27; DB 5; Length 10;
Best Local Similarity 83.3%; Pred. No. 0.94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMMMR 6
|:||||
Db 1 MLMMMR 6

RESULT 3
US-09-252-991A-29802
; Sequence 29802, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 29802
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29802

Query Match 79.4%; Score 27; DB 4; Length 61;
Best Local Similarity 71.4%; Pred. No. 5.7;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMMMR 7
|:||||
Db 3 MLMMMR 9

RESULT 4
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/745,206A
; FILING DATE: 19910815
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2161 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-745-206A-2

Query Match 79.4%; Score 27; DB 1; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLMMMR 6
|:||||
Db 3 MLMMMR 8

RESULT 5
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven

APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: May 31, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2161 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-49

Query Match 79.4%; Score 27; DB 1; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R 6
Db 3 M M M M M K 8

RESULT 6

US-08-455-543A-51
Sequence 51, Application US/084555543A
Patent No. 5792846
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: May 31, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 2161 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-51

Query Match 79.4%; Score 27; DB 1; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M M M M M R 6
Db 3 M M M M M K 8

RESULT 7

US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/745,206
FILING DATE: April 10, 1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/620,250
FILING DATE: 15-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELEPHONE: (619)238-0999

TELEFAX:

(619)238-0062

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 2161 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-08-223-305C-49

Query Match

79.4%; Score 27; DB 2; Length 2161;

Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M M M M M R 6
Db 3 M M M M M K 8

RESULT 8

US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 4, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELEPHONE: (619)238-0999

TELEFAX:

(619)238-0062

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:

LENGTH: 2161 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-08-223-305C-51

Query Match 79.4%; Score 27; DB 2; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
|||||:
Db 3 MMTMMK 8

RESULT 9

US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,363
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-51506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-311-363-2

Query Match 79.4%; Score 27; DB 2; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
|||||:
Db 3 MMTMMK 8

RESULT 10

US-08-891-254-3
; Sequence 3, Application US/08891254
; Patent No. 5776889
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min

; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: Hypersensitive Response
; TITLE OF INVENTION: Induced Resistance In Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,254
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/475,775
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 14603/10050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-891-254-3

Query Match 76.5%; Score 26; DB 1; Length 385;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
|||||:
Db 63 MMTMMSM 69

RESULT 11

US-08-819-539-3
; Sequence 3, Application US/08819539
; Patent No. 5859324
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: Hypersensitive Response
; TITLE OF INVENTION: Induced Resistance In Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,539
; FILING DATE: 17-MAR-1997

CLASSIFICATION: 800
PRIOR APPLICATION DATA: 08/475,775
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 14603/10050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1500
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-819-539-3

Query Match 76.5%; Score 26; DB 2; Length 385;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
Db 63 M M M M M S M 69

RESULT 12
PCT-US93-06243-2
Sequence 2, Application PC/TUS9306243
GENERAL INFORMATION:
APPLICANT: Zhong-Min Wei, David W. Bauer, Steven V.
APPLICANT: Beer, Alan Collmer, Sheng-Yang He, and Ron J. Laby
TITLE OF INVENTION: Elicitor of the Hypersensitive Response in Plants
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06243
FILING DATE: 19930630
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 907,935
FILING DATE: 01-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: CRF D-1172
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-06243-2

Query Match 76.5%; Score 26; DB 5; Length 385;

US-09-901-187c-12.ra1

Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
Db 63 M M M M M S M 69

RESULT 13
PCT-US96-08819-3
Sequence 3, Application PC/TUS9608819
GENERAL INFORMATION:
APPLICANT: Cornell Research Foundation, Inc.
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
TITLE OF INVENTION: RESISTANCE IN PLANTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08819
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,775
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/10051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-08819-3

Query Match 76.5%; Score 26; DB 5; Length 385;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
Db 63 M M M M M S M 69

RESULT 14
US-08-200-724A-2
Sequence 2, Application US/08200724A
Patent No. 5849868
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: Bauer, David W.
APPLICANT: Beer, Steven V.
APPLICANT: Collmer, Alan
APPLICANT: He, Sheng-Yang
APPLICANT: Laby, Ron J.
TITLE OF INVENTION: ELICITOR OF THE HYPERSENSITIVE RESPONSE
TITLE OF INVENTION: IN PLANTS

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/200,724A
FILING DATE: 23-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/10030
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-200-724A-2

Query Match 76.5%; Score 26; DB 2; Length 403;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 M M M M M R L 7
| | | | | :
Db 63 M M M M M S M 69

RESULT 15
US-09-030-270A-3
Sequence 3, Application US/09030270A
Patent No. 5977060
GENERAL INFORMATION:
APPLICANT: Zitter, Thomas A.
APPLICANT: Wei, Zhong-Min
TITLE OF INVENTION: INSECT CONTROL WITH A
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,270A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,226
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-030-270A-3

Query Match 76.5%; Score 26; DB 2; Length 403;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 M M M M M R L 7
| | | | | :
Db 63 M M M M M S M 69

Search completed: February 11, 2004, 17:13:40
Job time : 11.4167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 M M M M M M R L 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	7	10	US-09-901-187B-12
2	32	94.1	103	9	US-09-864-761-44210
3	29	85.3	42	9	US-09-864-761-45051
4	28	82.4	32	9	US-09-864-761-34021
5	28	82.4	223	12	US-10-108-260A-4500
6	28	82.4	255	12	US-10-017-161-2180
7	28	82.4	255	12	US-10-292-798-1826
8	27	79.4	18	9	US-09-864-761-35316
9	27	79.4	18	9	US-09-864-761-38185
10	27	79.4	33	9	US-09-864-761-45071
11	27	79.4	111	9	US-09-864-761-43445
12	27	79.4	305	10	US-09-886-055-35
13	27	79.4	305	11	US-09-779-679-14
14	27	79.4	305	11	US-09-779-679-16
15	27	79.4	305	11	US-09-779-679-18

16	79.4	305	11	US-09-779-679-22	Sequence 22, Appl
17	79.4	305	11	US-09-804-291-35	Sequence 35, Appl
18	79.4	305	12	US-10-387-629-42	Sequence 42, Appl
19	79.4	427	12	US-10-374-780A-2722	Sequence 2722, Ap
20	79.4	2161	12	US-10-375-253-2	Sequence 2, Appl
21	79.4	2166	14	US-10-029-413A-4	Sequence 4, Appl
22	79.4	2181	12	US-10-411-010-29	Sequence 29, Appl
23	79.4	2181	14	US-10-029-413A-18	Sequence 18, Appl
24	79.4	2181	14	US-10-029-413A-20	Sequence 20, Appl
25	79.4	2181	15	US-10-205-823-50	Sequence 50, Appl
26	76.5	18	9	US-09-864-761-35295	Sequence 35295, A
27	76.5	20	9	US-09-864-761-33376	Sequence 33376, A
28	76.5	22	9	US-09-864-761-40486	Sequence 40486, A
29	76.5	25	9	US-09-864-761-33850	Sequence 33850, A
30	76.5	26	9	US-09-864-761-42400	Sequence 42400, A
31	76.5	27	9	US-09-864-761-40516	Sequence 40516, A
32	76.5	27	9	US-09-864-761-41279	Sequence 41279, A
33	76.5	33	9	US-09-864-761-48633	Sequence 48633, A
34	76.5	50	9	US-09-864-761-44258	Sequence 44258, A
35	76.5	60	12	US-10-029-386-31671	Sequence 31671, A
36	76.5	69	9	US-09-864-761-42285	Sequence 42285, A
37	76.5	403	9	US-09-086-118-23	Sequence 23, Appl
38	76.5	403	9	US-09-835-684-3	Sequence 3, Appl
39	76.5	403	9	US-09-880-371-3	Sequence 3, Appl
40	76.5	403	9	US-09-879-248-3	Sequence 3, Appl
41	76.5	403	9	US-09-770-693-3	Sequence 3, Appl
42	76.5	403	10	US-09-766-348-3	Sequence 3, Appl
43	76.5	403	12	US-10-387-806-23	Sequence 23, Appl
44	76.5	403	12	US-10-441-736-3	Sequence 3, Appl
45	76.5	403	15	US-10-034-158-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-901-187B-12
; Sequence 12, Application US/09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Woloizin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Micheal S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment
; TITLE OF INVENTION: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-12

Query Match 100.0%; Score 34; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M M M M M R L 7
| | | | |
Db 1 M M M M M R L 7

RESULT 2
US-09-864-761-44210
; Sequence 44210, Application US/09864761

Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US 09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 44210
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005668.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
US-09-864-761-44210

Query Match 94.1%; Score 32; DB 9; Length 103;
Best Local Similarity 85.7%; Pred. No. 3.6;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
Db 16 M M M M M R M 22

RESULT 3

US-09-864-761-45051
Sequence 45051, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US 09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 45051
LENGTH: 42
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009933.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.67
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.69
US-09-864-761-45051

Query Match 85.3%; Score 29; DB 9; Length 42;
Best Local Similarity 71.4%; Pred. No. 7.1;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
Db 20 M M M M M K M 26

RESULT 4
US-09-864-761-34021
; Sequence 34021, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34021
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005100.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.74
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6
US-09-864-761-34021

Query Match 82.4%; Score 28; DB 9; Length 32;
Best Local Similarity 85.7%; Pred. No. 9.2;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMMMRL 7
Db 3 MMMMML 9
RESULT 5
US-10-108-260A-4500
; Sequence 4500, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4500
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4500

Query Match 82.4%; Score 28; DB 12; Length 223;
Best Local Similarity 85.7%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMMMRL 7
Db 1 MMMMML 7
RESULT 6
US-10-017-161-2180
; Sequence 2180, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2180
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-2180

Query Match 82.4%; Score 28; DB 12; Length 255;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMRL 7
Db 111 MMMVM 117
RESULT 7
US-10-292-798-1826
; Sequence 1826, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI


```

; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1826
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1826

Query Match      82.4%; Score 28; DB 12; Length 255;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MMTMMRL 7
Db      111 MMTMMVRM 117

RESULT 8
US-09-864-761-35316
; Sequence 35316, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30

; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1826
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1826

Query Match      82.4%; Score 28; DB 12; Length 255;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MMTMMRL 7
Db      111 MMTMMVRM 117

RESULT 9
US-09-864-761-38185
; Sequence 38185, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
```

```

; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35316
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035704.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
US-09-864-761-35316

Query Match      79.4%; Score 27; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 MMTMMMR 6
Db      13 MMTMMMK 18

RESULT 9
US-09-864-761-38185
; Sequence 38185, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 38185
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL035704.9
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
US-09-864-761-38185

Query Match 79.4%; Score 27; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
| | | | | | | |
Db 13 M M M M M M K 18

RESULT 10

US-09-864-761-45071
; Sequence 45071, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663

;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 45071
;; LENGTH: 33
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC025644.2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.84
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.59
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.76
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.57
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.57
US-09-864-761-45071

Query Match 79.4%; Score 27; DB 9; Length 33;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
| | | | | | | |
Db 10 M M M M M M K 15

RESULT 11

US-09-864-761-43445
; Sequence 43445, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30

;; PRIORITY APPLICATION NUMBER: PCT/US01/00663
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00662
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00661
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00670
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: US 60/234,687
;; PRIORITY FILING DATE: 2000-09-21
;; PRIORITY APPLICATION NUMBER: US 09/608,408
;; PRIORITY FILING DATE: 2000-06-30
;; PRIORITY APPLICATION NUMBER: US 09/774,203
;; PRIORITY FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 43445
;; LENGTH: 111
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006097.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.62
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.81
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
;; OTHER INFORMATION: SWISSPROT HIT: Q10699, EVALUATE 5.90e-01
US-09-864-761-43445

Query Match 79.4%; Score 27; DB 9; Length 111;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMR 6
|:|:|:|:
Db 1 MMMMMK 6

RESULT 12

US-09-886-055-35
;; Sequence 35, Application US/09886055
;; Patent No. US20020132273A1
;; GENERAL INFORMATION:
;; APPLICANT: STRYER, LUBERT
;; APPLICANT: ZOZULYA, SERGEY
;; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
;; FILE REFERENCE: 078003-0277150
;; CURRENT APPLICATION NUMBER: US/09/886,055
;; CURRENT FILING DATE: 2001-06-22
;; PRIORITY APPLICATION NUMBER: 60/213,812
;; PRIORITY FILING DATE: 2000-06-22
;; NUMBER OF SEQ ID NOS: 522
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 35
;; LENGTH: 305
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-886-055-35

Query Match 79.4%; Score 27; DB 10; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMR 7
|:|:|:|:
Db 45 MIMLMR 51

RESULT 13

US-09-779-679-14

;; Sequence 14, Application US/09779679
;; Publication No. US20030082757A1
;; GENERAL INFORMATION:
;; APPLICANT: Taupier, Raymond J
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Padigaru, Muralidhara
;; APPLICANT: Tchernev, Velizar T
;; APPLICANT: Mishra, Vishnu S
;; APPLICANT: Casman, Stacie
;; APPLICANT: Ballinger, Robert
;; APPLICANT: Vernet, Corine A
;; APPLICANT: Li, Li
;; APPLICANT: Spytek, Kimberly A
;; APPLICANT: Andrew, David P
;; APPLICANT: Mezes, Peter S
;; TITLE OF INVENTION: No. US20030082757A1 Proteins and Nucleic Acids Encoding the Sa
;; FILE REFERENCE: 15966-661
;; CURRENT APPLICATION NUMBER: US/09/779,679
;; CURRENT FILING DATE: 2001-02-08
;; PRIORITY APPLICATION NUMBER: USSN 60/181045
;; PRIORITY FILING DATE: 2000-02-08
;; PRIORITY APPLICATION NUMBER: USSN 60/183191
;; PRIORITY FILING DATE: 2000-02-17
;; PRIORITY APPLICATION NUMBER: USSN 60/180,929
;; PRIORITY FILING DATE: 2000-02-08
;; PRIORITY APPLICATION NUMBER: USSN 60/219758
;; PRIORITY FILING DATE: 2000-07-20
;; PRIORITY APPLICATION NUMBER: USSN 60/181339
;; PRIORITY FILING DATE: 2000-02-09
;; PRIORITY APPLICATION NUMBER: USSN 60/181344
;; PRIORITY FILING DATE: 2000-02-09
;; PRIORITY APPLICATION NUMBER: USSN 60/221341
;; PRIORITY FILING DATE: 2000-07-26
;; PRIORITY APPLICATION NUMBER: USSN 60/181392
;; PRIORITY FILING DATE: 2000-02-09
;; PRIORITY APPLICATION NUMBER: USSN 60/219585
;; PRIORITY FILING DATE: 2000-07-20
;; PRIORITY APPLICATION NUMBER: USSN 60/181157
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 14
;; LENGTH: 305
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-779-679-14

Query Match 79.4%; Score 27; DB 11; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMR 7
|:|:|:|:
Db 45 MIMLMR 51

RESULT 14

US-09-779-679-16
;; Sequence 16, Application US/09779679
;; Publication No. US20030082757A1
;; GENERAL INFORMATION:
;; APPLICANT: Taupier, Raymond J
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Padigaru, Muralidhara
;; APPLICANT: Tchernev, Velizar T
;; APPLICANT: Mishra, Vishnu S
;; APPLICANT: Casman, Stacie
;; APPLICANT: Ballinger, Robert
;; APPLICANT: Vernet, Corine A
;; APPLICANT: Li, Li
;; APPLICANT: Spytek, Kimberly A
;; APPLICANT: Andrew, David P
;; APPLICANT: Mezes, Peter S

```

; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-679-16

```

```

Query Match      79.4%; Score 27; DB 11; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MMTMMRL 7
Db      45 MIMLRL 51

```

RESULT 15

```

US-09-779-679-18
; Sequence 18, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344

```

```

; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 305.
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-679-18

Query Match      79.4%; Score 27; DB 11; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MMTMMRL 7
Db      45 MIMLRL 51

```

Search completed: February 11, 2004, 17:54:14
Job time : 25.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 MWMWRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 19Jun03: *
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	7	AAE14557	Human alpha-synuclein
2	32	94.1	103	ABG51419	Human liver peptid
3	32	94.1	103	ABG51419	Peptide #7431 enco
4	32	94.1	103	ABG51419	Human brain expres
5	32	94.1	103	ABG51419	Human bone marrow
6	32	94.1	103	ABG51419	Peptide #7581 enco
7	32	94.1	103	ABG51419	Human peptid enco
8	30	88.2	167	AAE14557	Arabidopsis thalia
9	30	88.2	193	AAE14557	Arabidopsis thalia

10	30	88.2	194	21	AAE14557	Arabidopsis thalia
11	30	88.2	212	21	AAE14557	Arabidopsis thalia
12	30	88.2	212	21	AAE14557	Arabidopsis thalia
13	30	88.2	214	21	AAE14557	Arabidopsis thalia
14	30	88.2	233	21	AAE14557	Arabidopsis thalia
15	30	88.2	235	21	AAE14557	Arabidopsis thalia
16	30	88.2	249	21	AAE14557	Arabidopsis thalia
17	30	88.2	249	21	AAE14557	Arabidopsis thalia
18	30	88.2	314	21	AAE14557	Arabidopsis thalia
19	30	88.2	334	21	AAE14557	Arabidopsis thalia
20	30	88.2	339	21	AAE14557	Arabidopsis thalia
21	30	88.2	344	22	AAE14557	Human polyptide
22	30	88.2	352	22	AAE14557	Novel human diago
23	30	88.2	353	21	AAE14557	Arabidopsis thalia
24	30	88.2	358	21	AAE14557	Arabidopsis thalia
25	29	85.3	42	22	AAE14557	Human liver peptid
26	29	85.3	42	22	AAE14557	Peptide #9086 enco
27	29	85.3	42	22	AAE14557	Human brain expres
28	29	85.3	42	22	AAE14557	Human bone marrow
29	29	85.3	42	22	AAE14557	Peptide #9409 enco
30	28	82.4	32	22	AAE14557	Human liver peptid
31	28	82.4	32	22	AAE14557	Human peptid #738
32	28	82.4	32	22	AAE14557	Peptide #767 enco
33	28	82.4	32	22	AAE14557	Protein #722 enco
34	28	82.4	32	22	AAE14557	Human brain expres
35	28	82.4	32	22	AAE14557	Peptide #748 enco
36	28	82.4	32	22	AAE14557	Peptide #761 enco
37	28	82.4	32	22	AAE14557	Peptide #725 enco
38	28	82.4	32	23	AAE14557	Human peptid enco
39	28	82.4	35	24	AAE14557	Human CEA exon #10
40	28	82.4	107	21	AAE14557	Arabidopsis thalia
41	28	82.4	120	21	AAE14557	Arabidopsis thalia
42	28	82.4	129	22	AAE14557	Drosophila melanog
43	28	82.4	1033	24	AAE14557	Human CEA protein
44	27	79.4	10	16	AAE14557	GAP-43 protein der
45	27	79.4	18	22	AAE14557	Human liver peptid

ALIGNMENTS

RESULT 1
AAE14557
ID AAE14557 standard; peptide; 7 AA.
XX AAE14557;
AC AAE14557;
DT 17-MAY-2002 (first entry)
XX Human alpha-synuclein aggregation inhibitor #12.
DE Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
KW multiple system atrophy; Hallervorden-Spatz disease; human.
XX Homo sapiens.
OS WO200204482-A1.
XX 17-JAN-2002.
XX 06-JUL-2001; 2001WO-US21379.
XX 07-JUL-2000; 2000US-217319P.
XX 28-MAR-2001; 2001US-279199P.
XX (PANA-) PANACEA PHARM INC.
XX Wolozin B, Ostretova-Golts N, Lebowitz MS;
XX WPI; 2002-179695/23.
XX Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper
XX Claim 40; Page 37; 52pp; English.
XX The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to the NAC (non-amyloid-beta protein component)
XX portion of human alpha-synuclein and inhibits its aggregation.
SQ Sequence 7 AA;
Query Match 100.0%; Score 34; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMMMRL 7
DB 1 MMMMRL 7
RESULT 2
ABG55058
ID ABG55058 standard; Peptide; 103 AA.
XX AC ABG55058;
XX DT 25-FEB-2003 (first entry)
DE Human liver peptide, SEQ ID No 33706.
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00664.
XX PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-48898/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX Claim 27; SEQ ID No 33706; 658pp; English.
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG59930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 103 AA;
Query Match 94.1%; Score 32; DB 22; Length 103;
Best Local Similarity 85.7%; Pred. No. 1.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMMMRL 7
DB 16 MMMMRL 22
RESULT 3
ABB39925
ID ABB39925 standard; Peptide; 103 AA.
XX AC ABB39925;
XX DT 04-FEB-2002 (first entry)
XX DE Peptide #7431 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00669.
XX PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human fetal liver -
XX Claim 27; SEQ ID No 32560; 639pp + sequence listing; English.
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 103 AA;
Query Match 94.1%; Score 32; DB 22; Length 103;
Best Local Similarity 85.7%; Pred. No. 1.7;

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -

XX Claim 27; SEQ ID No 33813; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.

XX Sequence 103 AA;

Query Match 94.1%; Score 32; DB 22; Length 103;
Best Local Similarity 85.7%; Pred. No. 1.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
Db 16 MMTMMRM 22

RESULT 7
ABG43195
ID ABG43195 standard; Peptide; 103 AA.

AC ABG43195;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 32860.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -

XX Claim 27; SEQ ID No 32860; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 103 AA;

Query Match 94.1%; Score 32; DB 23; Length 103;
Best Local Similarity 85.7%; Pred. No. 1.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMRL 7
Db 16 MMTMMRM 22

RESULT 8

AAG24998

ID AAG24998 standard; Protein; 167 AA.

XX AAG24998;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 28884.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 167;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNNMNR 6
 Db 92 MNNMNR 97

RESULT 9
 AAG51419
 ID AAG51419 standard; Protein; 193 AA.

XX AC AAG51419;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 65257.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 193;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
|||
Db 1 M M M M M M R 6

RESULT 10
AAG51418
ID AAG51418 standard; Protein; 194 AA.

XX AAG51418;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65256.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2

XX 06-SEP-2000

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 194;

Best Local Similarity 100.0%; Pred. No. 9.7; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
D5 2 M M M M M R 7

RESULT 11
AAG24997
ID AAG24997 standard; Protein; 212 AA.

AC AAG24997;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28883.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS EP1033405-A2.

PN 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

PF 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 212;
Best Local Similarity 100.0%; Pred.No.11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R 6
Db 137 M M M M M R 142

RESULT 12
AAG51694
ID AAG51694 standard; Protein; 212 AA.

XX AAG51694;
XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65635.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

```
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      88.2%; Score 30; DB 21; Length 212;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 M M M M M R 6
Db      137 M M M M M R 142

RESULT 13
AAG23345
ID AAG23345 standard; Protein; 214 AA.
XX
AC AAG23345;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26618.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134221.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
```

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 214;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTMMR 6
| | | | |
Db 2 MMTMMR 7

RESULT 14
AAG51436
ID AAG51436 standard; Protein; 233 AA.

XX AAG51436;
AC AAG51436;
XX 18-OCT-2000 (first entry)
DT
XX
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65281.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.


```
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      88.2%; Score 30; DB 21; Length 233;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMTMMR 6
Db 2 MMTMMR 7

RESULT 15
AAG51693
ID AAG51693 standard; Protein; 235 AA.
XX
AC AAG51693;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65634.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
```

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 235;
Best Local Similarity 100.0%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 M M M M M R 6
Db 160 M M M M M R 165

Search completed: February 11, 2004, 17:03:05
Job time : 33.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-5
Perfect score: 38
Sequence: 1 THRLPSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		A Geneseq 19Jun03:*	
1:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*		
2:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*		
3:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*		
4:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*		
5:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*		
6:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*		
7:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*		
8:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*		
9:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*		
10:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*		
11:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*		
12:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*		
13:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*		
14:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*		
15:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*		
16:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*		
17:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*		
18:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*		
19:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*		
20:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*		
21:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*		
22:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*		
23:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*		
24:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	7	20	AAV31365
2	38	100.0	7	23	AAE14550
3	38	100.0	1608	17	AAR98619
4	38	100.0	1711	17	AAR98605
5	35	92.1	163	22	AAU40892
6	34	89.5	356	22	ABG02462
7	33	86.8	9	16	AAR79577
8	33	86.8	9	21	AAV57976
9	33	86.8	18	22	ABG48096

10	33	86.8	18	22	ABB28073	Human peptide #724
11	33	86.8	18	22	ABB33247	Peptide #753 encod
12	33	86.8	18	22	ABB18710	Protein #709 encod
13	33	86.8	18	22	AAM54040	Human brain expres
14	33	86.8	18	22	AAM66430	Human bone marrow
15	33	86.8	18	22	AAM14300	Peptide #734 encod
16	33	86.8	18	22	AAM26710	Peptide #747 encod
17	33	86.8	18	22	AAM02029	Peptide #711 encod
18	33	86.8	18	23	ABG36079	Human peptide enco
19	33	86.8	25	22	AAG22141	Human cardiovascular
20	33	86.8	56	22	AAU54781	Propionibacterium
21	33	86.8	99	21	ABG32032	Human secreted pro
22	33	86.8	105	22	AAM23902	Human EST encoded
23	33	86.8	137	22	AAU49649	Propionibacterium
24	33	86.8	251	21	AAU44450	Mutant chicken C-S
25	33	86.8	533	14	AAR39705	Chicken pp60 c-src
26	33	86.8	533	21	AAU44447	Wild-type chicken
27	33	86.8	533	21	AAU44449	Mutant chicken C-S
28	33	86.8	533	21	AAU44451	Mutant chicken C-S
29	33	86.8	533	22	ABG44661	Amino acid sequenc
30	32	84.2	55	21	AAG57799	Arabidopsis thalia
31	31	81.6	57	22	AAU40356	Propionibacterium
32	31	81.6	71	23	ABP33576	Human DNA/RNA bind
33	31	81.6	86	22	AAU51556	Propionibacterium
34	31	81.6	135	22	AAU58876	Propionibacterium
35	31	81.6	168	22	AAU57083	Propionibacterium
36	31	81.6	168	22	ABG02649	Novel human diagno
37	31	81.6	202	22	ABG74608	Human hPPG-1 prote
38	31	81.6	231	22	AAU51839	Propionibacterium
39	31	81.6	321	22	AAG65582	Human RTA-like G P
40	31	81.6	321	22	AAE06768	Human G-protein co
41	31	81.6	321	22	AAU04366	Human G-protein co
42	31	81.6	321	22	AAG64124	Human G-protein-co
43	31	81.6	321	22	AAU19292	Human G-protein-co
44	31	81.6	321	23	ABP95616	Human GPCR polypep
45	31	81.6	321	23	ABG79184	Human mas proto-on

ALIGNMENTS

RESULT 1
AAV31365
ID AAV31365 standard; peptide; 7 AA.

XX AAV31365;

XX 01-NOV-1999 (first entry)

DE B. cereus T spore tight-binding peptide.

XX Bacterial spore; biopanning; phage-display library; Bacillus; vaccine;
KW pathogen detection; biological warfare agent; B. anthracis.

XX Bacillus cereus.

XX WO9936081-A1.

XX 22-JUL-1999.

XX 14-JAN-1999; 99WO-US00771.

XX 14-JAN-1998; 98US-0071411.

XX (UABR-) UAB RES FOUND.

XX Turnbough CL;

XX WPI; 1999-468943/39.

XX N-PSDB; AAX99291.

PT Identifying peptides that bind to the surface of bacterial spores by
biopanning phage-display library, useful as vaccines and diagnostic

PT agents
XX Disclosure; Page 13; 23pp; English.
XX
CC The invention provides peptides that bind to the surface of bacterial
CC spores. These peptides are identified by biopanning a phage-display
CC library with the spores. The library is incubated with spores and any
CC phage-spore complexes formed recovered by centrifuging. They are washed
CC thoroughly, then phage eluted with buffer, the eluate neutralised and
CC eluted phage amplified. The procedure is repeated, for 3 or 4 rounds of
CC panning, then individual clones purified, amplified and genomic DNA
CC extracted for determination of peptide-encoding sequences. Peptides
CC encoded by these sequences are then tested for ability to bind to target
CC spores. The peptides are used for capture and identification of
CC bacterial spores, particularly of the genus Bacillus, particularly for
CC detecting pathogens, or otherwise harmful species, in the environment
CC (e.g. air, water or food) or in clinical samples. They are also used to
CC protect against disease-causing spores (e.g. by incorporation in
CC protective masks), and in vaccines to generate a protective antibody
CC response. The peptides-specific antibodies are useful as detection
CC reagents, e.g. in enzyme-linked immunosorbent assay. The peptides are
CC species specific, i.e. they can differentiate between the possible
CC biological warfare agent B. anthracis and related species, widely present
CC in the environment and likely to give rise to false positive results.
CC Sequences AAY31365-372 represent amino acid sequences of B. cereus T
CC spore tight-binding peptides.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 THRLPSR 7
|||||||
Db 1 THRLPSR 7

RESULT 2
AAE14550
ID AAE14550 standard; peptide; 7 AA.
XX
AC AAE14550;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human alpha-synuclein aggregation inhibitor #5.
XX
KW Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
KW multiple system atrophy; Hallervorden-Spatz disease; human.
XX
OS Homo sapiens.
XX
PN WO200204482-A1.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21379.
XX
PR 07-JUL-2000; 2000US-217319P.
PR 28-MAR-2001; 2001US-279199P.
XX
PA (PANA-) PANACEA PHARM INC.
XX
XX Wolozin B, Ostrétova-Golts N, Lebowitz MS;
PI
XX WPI; 2002-179695/23.
XX
XX Determination of an agent capable of inhibiting aggregation of alpha
PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper

XX Claim 40; Page 37; 52pp; English.
PS
XX
CC The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to C-terminal portion of human alpha-synuclein and
CC inhibits its aggregation.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 THRLPSR 7
|||||||
Db 1 THRLPSR 7

RESULT 3
AAR98619
ID AAR98619 standard; Protein; 1608 AA.
XX
AC AAR98619;
XX
DT 10-DEC-1996 (first entry)
XX
DE Borna disease virus polymerase.
XX
KW Borna disease virus; BDV; G-protein; p57; nervous system disease;
KW neuro-psychiatric disease; schizophrenia; diagnosis; therapy;
KW vaccine; antibody.
XX
OS Borna disease virus strain V.
XX
PN WO9621020-A2.
XX
PD 11-JUL-1996.
XX
PF 05-JAN-1996; 96WO-US00418.
XX
PR 04-JAN-1996; 96US-0582776.
PR 06-JAN-1995; 95US-0369822.
PR 04-MAY-1995; 95US-0434831.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Briese T, Kliche S, Lipkin WI, Schneemann A, Schneider PA;
PI Stitz L;
XX
DR WPI; 1996-333995/33.
DR N-PSDB; AAT38104.
XX
PT Borna disease virus (BDV) nucleotide and protein sequences - useful
PT for the diagnosis and treatment of infection and non-BDV related
PT neuro-logic and neuro-psychiatric disease
XX
PS Claim 2; Fig 2; 186pp; English.
XX
CC Borna disease virus (BDV) polymerase (AAR98619), or pol or p180, was
CC identified from an ORF on the virus genome strand (AAT38104).
CC The amino acid sequence for pol after splice modification is given
CC in AAR98605. Recombinant pol (recpol) can be expressed in transformed
CC host (partic. mammalian) cells. It is useful in assays for
CC detecting BDV infection and for diagnosing non-BDV related neurologic
CC and neuropsychiatric diseases. It may also be incorporated into
CC vaccines and used to raise anti-BDV antibodies.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-5
Perfect score: 38
Sequence: 1 THRLPSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	38	100.0	7	10	US-09-901-187B-5
2	38	100.0	7	11	US-09-229-751A-72
3	33	86.8	18	9	US-09-864-761-34008
4	33	86.8	25	9	US-09-764-869-915
5	33	86.8	25	12	US-10-227-577-915
6	33	86.8	25	15	US-10-091-504-915
7	33	86.8	362	12	US-10-369-493-10705
8	32	84.2	330	12	US-10-369-493-10663
9	32	84.2	542	12	US-10-369-493-12602
10	31	81.6	71	12	US-09-864-408A-5098
11	31	81.6	321	10	US-09-920-068A-2
12	31	81.6	321	12	US-10-017-161-1050
13	31	81.6	321	12	US-10-239-421-2
14	31	81.6	321	12	US-10-258-768-1
15	31	81.6	321	12	US-10-088-726-19

16	31	81.6	321	12	US-10-305-555-4	Sequence 4, Appli
17	31	81.6	321	12	US-10-321-807-10	Sequence 10, Appl
18	31	81.6	321	12	US-10-237-467-8	Sequence 8, Appli
19	31	81.6	321	12	US-09-801-944B-268	Sequence 268, App
20	31	81.6	321	15	US-10-188-405-2	Sequence 2, Appli
21	31	81.6	321	15	US-10-183-116-35	Sequence 35, Appl
22	31	81.6	321	15	US-10-079-384-8	Sequence 8, Appli
23	31	81.6	347	12	US-10-292-798-892	Sequence 892, App
24	31	81.6	504	12	US-10-289-762-1110	Sequence 1110, Ap
25	30	78.9	43	12	US-10-029-386-27495	Sequence 27495, A
26	30	78.9	72	10	US-09-764-847-634	Sequence 634, App
27	30	78.9	72	15	US-10-092-154-634	Sequence 634, App
28	30	78.9	85	12	US-09-833-245-969	Sequence 969, App
29	30	78.9	117	15	US-10-156-761-10633	Sequence 10633, A
30	30	78.9	186	12	US-10-158-057-223	Sequence 223, App
31	30	78.9	289	12	US-10-094-749-2399	Sequence 2399, App
32	30	78.9	301	9	US-09-925-299-786	Sequence 786, App
33	30	78.9	301	11	US-09-925-299-786	Sequence 786, App
34	30	78.9	301	15	US-10-106-698-4845	Sequence 4845, Ap
35	30	78.9	389	10	US-09-738-626-4492	Sequence 4492, Ap
36	30	78.9	429	15	US-10-156-761-13110	Sequence 13110, A
37	30	78.9	778	15	US-10-266-388-2	Sequence 2, Appli
38	30	78.9	797	15	US-10-266-388-4	Sequence 4, Appli
39	30	78.9	842	11	US-09-291-417-91	Sequence 91, Appli
40	30	78.9	846	15	US-10-303-683-22	Sequence 22, Appl
41	30	78.9	1066	12	US-10-369-493-22668	Sequence 22668, A
42	30	78.9	1702	10	US-09-738-973-434	Sequence 434, App
43	30	78.9	1702	10	US-09-854-133-434	Sequence 434, App
44	30	78.9	1702	15	US-10-144-649A-434	Sequence 434, App
45	30	78.9	6396	12	US-09-940-316B-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1

US-09-901-187B-5
; Sequence 5, Application US 09901187B
; Patent No. US20020151454A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Micheal S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatme
; TITLE OF INVENTION: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-5

Query Match 100.0%; Score 38; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THRLPSR 7
|||||
Db 1 THRLPSR 7

RESULT 2

US-09-229-751A-72
; Sequence 72, Application US/09229751A

Publication No. US20030044838A1
GENERAL INFORMATION:
APPLICANT: Turnbough, Charles K
TITLE OF INVENTION: PEPTIDE LIGANDS THAT BIND TO SURFACES OF BACTERIAL CELLS
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: Glenna Hendricks
STREET: P.O. Box 2509
CITY: Fairfax
STATE: VA
COUNTRY: USA
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/229,751A
FILING DATE: 14-Jan-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna M.
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: turn
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 425-8405
TELEFAX: (703) 425-8406
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-229-751A-72
Query Match 100.0%; Score 38; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 THRLPSR 7
Db 1 THRLPSR 7
RESULT 3
US-09-864-761-34008
Sequence 34008, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34008
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006017.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.9
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
OTHER INFORMATION: EST_HUMAN HIT: AW997046.1, EVALUE 2.00e-04
US-09-864-761-34008
Query Match 86.8%; Score 33; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 THRLPS 6
Db 11 THRLPS 16
RESULT 4
US-09-764-869-915
Sequence 915, Application US/09764869
Patent No. US20020061521A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764,869
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 915
LENGTH: 25
TYPE: PRT